

```
Db 945 TTCAAGGATTTCTCTCCCTCAGACCTCCTAGTATAGTGGATTAACAGGCAACCCGACACAC 1004
Qy 2872 GCCGTGATTAATTTTTTTGTAATTTTATAGACAGAGGTTTACACATGTTGGCCAGGCTG 2931
Db 1005 GCCGTGATTAATTTTTTTGTAATTTTATAGACAGAGGTTTACACATGTTGGCCAGGCTG 1064
Qy 2932 GTCTCGAATCCCTGACCTTAGGTGATACACCTGCTGACCTTCCCAAGTGTGGGATTA 2991
Db 1065 GTCTCGAATCCCTGACCTTAGGTGATACACCTGCTGACCTTCCCAAGTGTGGGATTA 1124
Qy 2992 TAGGATAGGACCATGGGCCAGTATTAAGAAAGTTAAAGGACATGGCAATGACAC 3051
Db 1125 TAGGATAGGACCATGGGCCAGTATTAAGAAAGTTAAAGGACATGGCAATGACAC 1184
Qy 3052 GCCATATCAGCTTTCCCTGCCAAAGCAAGGCAAGCTCTGGGCTACCTTCTGGGTT 3111
Db 1185 GCCATATCAGCTTTCCCTGCCAAAGCAAGGCAAGCTCTGGGCTACCTTCTGGGTT 1244
Qy 3112 TCTACTTCCAAAGGACATGACAACTGAGGCTTGGAGACCATTCATCCACTCT 3171
Db 1245 TCTACTTCCAAAGGACATGACAACTGAGGCTTGGAGACCATTCATCCACTCT 1304
Qy 3172 AGGCTCCCTATGGAGATGAGTCCAGACAGGGAAGGCTCTGACAGGCTGACCA 3231
Db 1305 AGGCTCCCTATGGAGATGAGTCCAGACAGGGAAGGCTCTGACAGGCTGACCA 1364
Qy 3232 GGGCTCTGATCCCTACCAACCCCAATCGGTCTCTCTACAGGACCCAGCCAC 3291
Db 1365 GGGCTCTGATCCCTACCAACCCCAATCGGTCTCTCTACAGGACCCAGCCAC 1424
Qy 3292 CTGCTGACAGCCCATGCTGGCCATGACATCACTTACCAAGCAAGTGGCTAATGCC 3351
Db 1425 CTGCTGACAGCCCATGCTGGCCATGACATCACTTACCAAGCAAGTGGCTAATGCC 1484
Qy 3352 GCTTAGGCTCTCTCCCGCGCTGCTGTGTGGGGGAGCACTTACAGCTGTAT 3411
Db 1485 GCTTAGGCTCTCTCCCGCGCTGCTGTGTGTGGGGGAGCACTTACAGCTGTAT 1544
Qy 3412 ATGGCAGTTCTTAATCTTCTGCTCTGCTACTACTATCATCTCGCTTATTTATAGTAA 3471
Db 1545 ATGGCAGTTCTTAATCTTCTGCTCTGCTACTACTATCATCTCGCTTATTTATAGTAA 1604
Qy 3472 GCTGACAGGCTGTGGCCGGGGGCTGTGGGAAGATGTGGCTGGGCTGGGAGCTGGAGC 3531
Db 1605 GCTGACAGGCTGTGGCCGGGGGCTGTGGGAAGATGTGGCTGGG-----GCTGGAGC 1657
Qy 3532 TCCCTGGGGCTCCAGCAGCTCAGGGCCAGTGCCAGCAGTCCACTTACAGCACTAAGCT 3591
Db 1658 TCCCTGGGGCTCCAGCAGCTCAGGGCCAGTGCCAGTCCACTTACAGCACTAAGCT 1717
Qy 3592 GGGCTCTGACAGCTCTGTGGGCACTGTAGAGTGTAGGCTGTGGGCTGGGGCTGGAGAG 3651
Db 1718 GAGCTCTGACAGCTGTGGGCACTGTAGAGTGTAGGCTGTGGGCTGGGGAGAG 1776
Qy 3652 TAAAGAGTCACTGAGAGGCTGTCAAGCAGGCAAGAGGTTTAAAGCACTCC 3711
Db 1777 TAAAGAGTCACTGAGAGGCTGTCAAGCAGGCAAGAGGTTTAAAGCACTCC 1836
Qy 3712 TCCAAACCCAGAGAGACCCCTGAGAGCCAGGCTTTGTCTGGCCCACTTACTGGCTGT 3771
Db 1837 TCCAAACCCAGAGAGACCCCTGAGAGCCAGGCTTTGTCTGGCCCACTTACTGGCTGT 1896
Qy 3772 TTTTACTGATCCCAAGAGATCATAGGCCCAATATATATTAAAAAAGAGAGAG 3831
Db 1897 TTTTACTGATCCCAAGAGATCATAGGCCCAATATATATTAAAAAAGAGAGAG 1956
Qy 3892 AGAAGAGAGAGAGATGAGTCTCACTGTGTGTCTCAGAGCTGTCTGAACCTCTAGG 3891
Db 1957 AGAAGAGAGAGAGATGAGTCTCACTGTGTGTCTCAGAGCTGTCTGAACCTCTAGG 2016
Qy 3892 CTCAAGCAATCCCTGCTTACCTCCCAAGGGGCTGGATTACAGTGTGAGCTCTG 3951
Db 3951 CTCAAGCAATCCCTGCTTACCTCCCAAGGGGCTGGATTACAGTGTGAGCTCTG 3951

Db 2017 CTCAGCAATCCCTGCTTACCTCCCAAGGGGCTGGATTACAGTGTGAGCTCTG 2076
Qy 3952 CACTTGACCAACACATGTAATTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGGTTTCC 4011
Db 2077 CACTTGACCAACACATGTAATTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGGTTTCC 2136
Qy 4012 TCCATCACAGGCTGAGTGTGAGGCAATCTTGCTCAGTCACTCTGCTCC 4071
Db 2137 TCCATCACAGGCTGAGTGTGAGGCAATCTTGCTCAGTCACTCTGCTCC 2196
Qy 4072 CAGTGCAGAGATTCCTGCTTACCTGCTGAGTACCTGATTAATAGCAGCAGC 4131
Db 2197 CAGTGCAGAGATTCCTGCTTACCTGCTGAGTACCTGATTAATAGCAGCAGC 2256
Qy 4132 ACCAGGCTGCTAATTTTTTTTTTTTTTTTTTTTTTTTGGATAGAGAGGTTTCATCAT 4191
Db 2257 ACCAGGCTGCTAATTTTTTTTTTTTTTTTTTTTTTTTGGATAGAGAGGTTTCATCAT 2216
Qy 4192 GTTGGCAGGCTGCTGTAACCCCTGACCTCAAGTATCCACCACTGCGCTCCCA 4251
Db 2317 GTTGGCAGGCTGCTGTAACCCCTGACCTCAAGTATCCACCACTGCGCTCCCA 2376
Qy 4252 AGTCTGGATTTACAGTGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4311
Db 2377 AGTCTGGATTTACAGTGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2436
Qy 4312 TATTTTAAATTAATGTTTATCTAAGGCAAGTACAGTACGCTGCTGTATGCCA 4371
Db 2437 TATTTTAAATTAATGTTTATCTAAGGCAAGTACAGTACGCTGCTGTATGCCA 2496
Qy 4372 GCATTTGAGGGGCAAGGTGCGGGATCACTGAGCTTGGAGTTACAGGTGGGCAACA 4431
Db 2497 GCATTTGAGGGGCAAGGTGCGGGATCACTGAGCTTGGAGTTACAGGTGGGCAACA 2556
Qy 4432 TAGTGAGACCCGCTCTACCAAAAATTTAAATAATAGCTGGAGTGTGGCTATTTGCC 4491
Db 2557 TAGTGAGACCCGCTCTACCAAAAATTTAAATAATAGCTGGAGTGTGGCTATTTGCC 2616
Qy 4492 TGTGCTCCACTCTTGTGGAGTGTGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAG 4551
Db 2617 TGTGCTCCACTCTTGTGGAGTGTGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAG 2676
Qy 4552 GCTGAGTACATGATCAACCACTGACCTGAGTGCAGCTATCTCAAAA 4611
Db 2677 GCTGAGTACATGATCAACCACTGACCTGAGTGCAGCTATCTCAAAA 2736
Qy 4612 GCAACCAATATATGTATCTAAACGTAAGTATTAATCAAGAAATATATGACTT 4671
Db 2737 GCAACCAATATATGTATCTAAACGTAAGTATTAATCAAGAAATATATGACTT 2796
Qy 4672 TTAATTTGAAAAGCATTAATGATTAATGATTTTAAATATCAATATGAAATCT 4731
Db 2797 TTAATTTGAAAAGCATTAATGATTAATGATTTTAAATATCAATATGAAATCT 2856
Qy 4732 TGTGTTCTTAATATGCTGCAACAAGGCACTTTGTTTACTAGGCAACAAGTAC 4791
Db 2857 TGTGTTCTTAATATGCTGCAACAAGGCACTTTGTTTACTAGGCAACAAGTAC 2916
Qy 4792 TTTAAAAAAGTTAGGCGCAGCCACAGGGGCTCAACCTGTATCCAGCACTTTGGAG 4851
Db 2917 TTTAAAAAAGTTAGGCGCAGCCACAGGGGCTCAACCTGTATCCAGCACTTTGGAG 2976
Qy 4852 GCCAAGGAGAGATCACTTGAAGCCAGAGTTTGAAGCTGAGCAACATAGGAGATC 4911
Db 2977 GCCAAGGAGAGATCACTTGAAGCCAGAGTTTGAAGCTGAGCAACATAGGAGATC 3036
Qy 4912 CTGATCTGTCTTATTAATAATTTAAATAATTTGCTAGGCTTTGCTTACACCGTAA 4971
Db 3037 CTGATCTGTCTTATTAATAATTTAAATAATTTGCTAGGCTTTGCTTACACCGTAA 3096
Qy 4972 TCCAGCACTTTGGAGGCGAGCGGCT 5000
Db 3097 TCCAGCACTTTGGAGGCGAGCGGCT 3125
```


XX WO200200928-A2.
 XX 03-JAN-2002.
 XX 02-JUL-2001; 2001WO-EP07537.
 XX 30-JUN-2000; 2000DE-1032529.
 XX 01-SEP-2000; 2000DE-1043826.
 XX (EPIC-) EPIGENOMICS AG.
 XX Olek A. Piepenbrock C, Berlin K;
 XX WPI; 2002-130909/17.
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 XX for diagnosis and treatment of diseases associated with abnormal
 XX cytosine methylation
 XX
 XX Claim 1; SEQ ID NO 2219; 32pp + Sequence Listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, aneuploidy, cancer, acute myeloid
 CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/intestinal bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 SO Sequence 6396 BP; 1584 A; 89 C; 1657 G; 3066 T; 0 other;

Query Match 25.7%; Score 1284.2; DB 24; Length 6396;
 Best Local Similarity 66.7%; Pred. No. 1.5e-313;
 Matches 2107; Conservative 0; Mismatches 694; Indels 357; Gaps 5;

QY 122 GCACCCACACGACGATACCTATGTAACCAACCTGCACCATGATACCTAT 181
 DB 3596 GTAAATTATATGATACGATATATATATATATATATATATATATATAT 3654
 QY 182 GTAAACACCTGCTGATCTGACACGATATCCACGAGCTTATAGATGAAAAAGTG 241
 DB 3655 GTAA-TTAAATTGTAATTTGTATACGATATTTA-GATTTTAAAGTAAAAAAGTG 3712
 QY 242 GTGTGTAGAAAAATCACTGCATCTCAGCATAGTAAACGCTATACATTTCAAGAGA 301
 DB 3713 GTGTGTAGAAAAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3772
 QY 302 GAGGTGACAGAAAGGAGATGAGAGTGGTTTAAACACACAGTGCATTTATTAAT 361
 DB 3773 GAGGTGATAGAAAGGAGATGAGAGTGGTTTAAACATTAAGATATTTATTAAT 3832
 QY 362 CAGGCTCTGGAAGTTAGTCCAAAAACACATCTCAATCCCTGAGTGTGGA 421
 DB 3833 TAGGCTTTTGGAAAGTTAGTTTAAATTTTAAATTTTAAATTTTGTAGTGTGA 3892
 QY 422 TTTAAATGCAACATCCCTAAGGCCACAGACTCAGACTGTGAGAAAGTCCAGAAACTG 481
 DB 3893 TTTAAATGATATATTTTAAAGTTATAGATTTTATTTTATTTTATTTTATTTT 3952
 QY 482 CCCGTTTAAATTAATTTGGCGATTTCTTACGGCTCTAAGACCAACCACTGCTGC 541
 DB 3953 TTGCTTTTAAATTAATTTGGCGATTTTACGGCTTTTAAAGATTAAGATTTATTTGT 4012
 QY 542 CTAGAGCTGCTCTCTTCAATTGAAACAATAGAGAGAGTGTAGTAAACCCACCA 601
 DB 4013 TTAGAGTTTGTGTTTTTTTATGATATATTAAGAGAGTGTAGTAAATTTATTTA 4072
 QY 602 CTTCACACAGCTTACGAGAGCCCTGATGATGATGATTAATTAATTTATTTGAATCA 661
 DB 4073 TTTTAAATGCTTACGAGAGCTTTTGTAGTATGATGATTAATTAATTTATTTGAATTA 4132

QY 662 CATGCTGATTTTCCACAGCTGCCCGTGGGATCTGGGCAATTAATCCATATATGCACT 721
 DB 4133 TATGTTAGATTTTATATATATGTTCTGCGGATTTGGGATTTATTTATTTATTTATTT 4192
 QY 722 GCGTGGCTGNAAGCCACAGCATTAACCTCCAGGCTCTGCTGCTACACCCACAGATC 781
 DB 4193 GGTGTTGGAGATTTATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4252
 QY 782 ACCCCCTCCACAGAGCCCGGAGGCTTCTCTTCCATCTCTGCAAGCAACTATGAT 841
 DB 4253 ATTTTATTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATTT 4312
 QY 842 GGGCCCTGCGCAACCAACAGCCAGATATGATATGATCAGACAGCCACCAATCAGAG 901
 DB 4313 GGGTTTGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4372
 QY 902 CTCTGCTGACATATGACAGATTTCTGCTATTTTACAGGCTGATGAATTTCCACCA 961
 DB 4373 TTTTGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4432
 QY 962 CACCATCTTTTCAATTAAGGCACTGAGCTGAGAGAGAGCTGAACCTACCGGGGT 1021
 DB 4433 TATTAATTTTATTTATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTA 4492
 QY 1022 CACACACACAGGTGCGAAGGCTGGGAGACCAAGAACAGACCTGACAGCCGGTA 1081
 DB 4493 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4552
 QY 1082 TTCAATTTCTTCCATAGCCACAGGCTGTCAAAAGACCCAGGCTAGTCAAGAGCTCT 1141
 DB 4553 TTTATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4612
 QY 1142 CCTCTCTGAGATGCTCTGCGCACAGAGTGAAGTCAACAGCCCTTAACCCCAAC 1201
 DB 4613 TTTTATTTGAGAGCTTTTGTATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 4672
 QY 1202 TCTCTCTGCAAGCCTCAAGGCTGCAAGACACAGTGTGAGACAGATCTTATAGCTCTG 1261
 DB 4673 TTTTATTTGAGGCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4732
 QY 1262 TTTAGGCGCATGATGAGGAGGCTGCTCCCTTAATTTCAAGCCCTGCTCAAGCCCAAC 1321
 DB 4733 TTTAGGCTTATGATGAGGAGGCTGCTCCCTTAATTTCAAGCCCTGCTCAAGCCCAAC 1381
 QY 1322 CCTCCAGAAATTAAGAGGCGCATGAGGCTGCTCAAGCCCTGCTCAAGCCCAAC 1381
 DB 4793 TTTTAAAGAAATTAAGAGGCGTATGATGATGATGATGATGATGATGATGATGATGATG 4852
 QY 1382 ATTCAAGAAAGGACTTAAGACAGACTCTCTTGTGAGAGTCTGAGCTTATGAGAGTCAAG 1441
 DB 4853 ATTCAAGAAAGGACTTAAGATTAAGATTTTCTGAGAGTCTGAGCTTATGAGAGTCAAG 4912
 QY 1442 TGAGGAGGCTGACACATCACTGCTGAGGAGGCAAGCCCTTAAGAGTGGGAGGCACTG 1501
 DB 4913 TGAGGAGGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4972
 QY 1502 GCCACAGAGTCCAGGAGGAGTCCACAGCTAGTCCGACAGACTCTGCTGAGGAGTCAAG 1561
 DB 4973 GTATATAGATTTAAGGAGTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTT 5032
 QY 1562 ACCCAGCTGAAACCCACCTGATGATTAAGAGTGTGAGGAGGAGGAGGAGGAGGAGG 1621
 DB 5033 ATTTATTTGGAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 5092
 QY 1622 GAGGCTTCTGAGGTTGATGAGGCTATGCTGCTTGTGAGGATTAATGCAACAGAGCA 1681
 DB 5093 GAGGTTTATTAAGGTTGATGAGGCTATTTTGTGCTATTTGAGGAGTATGATTAATTAAG 5152
 QY 1682 GCACATGCGCAGGCTGTGAGGCACTGTGTCTGTGCAAAATGCTGAGAGTGGAGTA 1741
 DB 5153 GATATATGCTAGGCTGTGAGGATTTGTGTGTTTGTGTAAGTGTGAGGAGGAGTA 5212
 QY 1742 GCTTGTGTGATCAAGAGAGCAAGCAGCAGTGTGCTGCAAGCAAAACACAGAGGA 1801

```
Db 5213 GTTTGGTGTATATAGG---TATAGTTAGTTAGTGTGTAGTAAATATATAGGAA 5269
Qy 1802 AGAATGAGGGGCGATCATCTACTGACAAAATATTATATAGAGCTCCCTAAAAAAG 1861
Db 5270 AGAATGAGGGGCGTAAATTA--- 5291
Qy 1862 AAGCTCTCTTTCATAGAGAGAGAGAGGGGGTTTGTCTTATTAATATAGG 1921
Db 5292 --- 5291
Qy 1922 AGGAGCCGCCCTCAAAAAATAGGAGGAGGAGGAGCCCAAGACCCGTTGGTGTGT 1981
Db 5292 --- 5291
Qy 1982 TTCAGGGGAGCTCGAACCTTTAGAGGAGCGTGGAGAACCGTGTATTCAGGCTCT 2041
Db 5292 --- 5291
Qy 2042 CGAGAGAAAAAGAGCGCGCCCAAAAAATATCCCTCCGGGCGATAGAAAATGTGCC 2101
Db 5292 --- 5291
Qy 2102 TCTTCAAAAAGATGAGAGAGAGCGGAGTTGTATGTGTATTTTAAACCCAG 2161
Db 5292 --- 5291
Qy 2162 GTAGNNNNNNNNNTGCTTCACTAAATTTTATGAGCGCTTACAGAAACACAGAG 2221
Db 5292 --- 5291
Qy 2222 GAGCTCCATCTGAGAGAGAAACAGCGAGAAACAGGAGATCCGTATATATTCAA 2281
Db 5339 GAGTTTATTTTGGAGAGAAATAGTGGAAATAGGATATTTTGTATATATTTAA 5398
Qy 2282 GTAGTATATGCTCTCTAGAAATATCAAGCAAGTGAGAGACACAGACCGGTGG 2341
Db 5399 GTAGTATATGCTCTCTAGAAATATCAAGCAAGTGAGAGACACAGACCGGTGG 2341
Qy 2342 CAGTGGGCTCTATTTCCAGGTTGGATGGTGGGAAACCTCTTAAAGGAGACCTGCA 2401
Db 5459 TAGTGGGCTTATTTTATTTAGTGGATGGTGGGAAATATTTTAAAGGAGAAATTTGA 5518
Qy 2402 GTGGAGAGAACCATGACAGATATCTCAGAGAGAGCTTCCAGGAGAGAGATCAGAG 2461
Db 5519 GTGGAGAGAAATATATAGTATTTAGAGAGATTTTATAGTATAGAGATATAGAG 5578
Qy 2462 GTGAGAGAGCCCTGAGGACCATTCATTAACATCATTTTGAATCTCTACAGCTAGG 2521
Db 5579 GTGAGAGAGCTTGGAGTATTTAGTAAATATATTTAGTATTTTATTTAGTAGG 5638
Qy 2522 TTCCATTTATGGGAATGGAATATGTTGAGAGAGGCTGCTCCCTCCCATCTTC 2581
Db 5639 TTTTATTTATGGGAATGGAATATGTTGAGAGAGGCTGCTTGTGTTTATTTATTTT 5698
Qy 2582 TCACACTAGGCTGTGAGAGAGCTTGGAGCTTAAAGAACAGATGGCTGAGAACACTG 2641
Db 5699 TTAATATAGGCTGTGAGAGAGCTTGGAGCTTAAAGAACAGATGGCTGAGAACACTG 2641
Qy 2642 CCTAGCCAGAGAGAGCTTGGAGCTTAAAGAACAGATGGCTGAGAACACTG 2641
Db 5759 TTTAGTTAGAGAGATTTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5818
Qy 2702 TGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2761
Db 5819 TGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2761
Qy 2762 CTTAGAGAGAGAGAGAGCTTGGAGCTTAAAGAACAGATGGCTGAGAACACTG 2821
Db 5879 TTTAGAGAGAGAGAGAGCTTGGAGCTTAAAGAACAGATGGCTGAGAACACTG 2821
Qy 2822 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2881
Db 5938 TTTAGAGAGAGAGAGAGCTTGGAGCTTAAAGAACAGATGGCTGAGAACACTG 2881

Db 5939 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5998
Qy 2882 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2941
Db 5999 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6058
Qy 2942 CCTGACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3001
Db 6059 TTTGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6118
Qy 3002 CCAGTGGGCGGAGTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3061
Db 6119 TTTATGCGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6178
Qy 3062 GTCTTCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3121
Db 6179 GTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6238
Qy 3122 AAGGAGTCTCAGACTGCGAGGCTTGGAGACCACTTCATCACCCTCTAGGCTCTCA 3181
Db 6239 AAGGAGTCTCAGACTGCGAGGCTTGGAGACCACTTCATCACCCTCTAGGCTCTCA 3181
Qy 3182 TGGAGAGTGTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3241
Db 6299 TGGAGAGTGTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3241
Qy 3242 TCCCTACAAACCCCATGCGTGTCTCTACACAG 3279
Db 6359 TTTTATTAATTTTATTAATCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6358

RESULT 7
AAS32727/c
AAS32727 standard; DNA; 17792 BP.
AAS32727;
17-DEC-2001 (first entry)

Human genomic DNA for novel endocrine antigen, SEQ ID No 681.
Human: endocrine antigen; ds; cytosolic; antihypertensive; antidiabetic;
thyroid-active; adrenal-active; androgenic; gastric; gene therapy;
antihypertensive; antidiabetic; endocrine disorder; hormone imbalance;
reproductive disorder; endocrine cancer; pancreatic disorder;
diabetes mellitus; adrenal gland disorder; hirsutism; thyroid disorder;
hypertension; hypothalamic disorder; vanishing testes syndrome.
Homo sapiens.
WO20015319-A2.
02-AUG-2001.
17-JAN-2001; 2001WO-US01335.
31-JAN-2000; 2000US-0179065.
04-FEB-2000; 2000US-0180628.
24-FEB-2000; 2000US-0184664.
02-MAR-2000; 2000US-0186350.
16-MAR-2000; 2000US-0189874.
17-MAR-2000; 2000US-0190076.
18-APR-2000; 2000US-0198123.
19-MAY-2000; 2000US-0205515.
07-JUN-2000; 2000US-0209467.
28-JUN-2000; 2000US-0214886.
30-JUN-2000; 2000US-0215135.
07-JUL-2000; 2000US-0216647.
07-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-0217487.
14-JUL-2000; 2000US-0217496.
14-JUL-2000; 2000US-0218280.
26-JUL-2000; 2000US-0220963.
```


hypothalamus and testes (e.g. vanishing testes syndrome), many examples of diseases and disorders are given in the specification. The present sequence is genomic DNA fragment form a gene encoding an endocrine antigen of the invention.
 Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIP0 at ftp.wip0.int/pub/published_pfc_sequences.

Sequence 17792 BP: 4372 A; 4626 C; 4413 G; 4381 T; 0 other;

Query Match 7.5%; Score 375.6; DB: 22; Length 17792; Best Local Similarity 64.3%; Pred. No. 7.9e-84; Matches 748; Conservative 0; Mismatches 359; Indels 57; Gaps 10;

3843 GAGAGGATGATCTCAGTGTGTCAGAGCTGCTGCAAGCTCTGAGTCAAGCAATC 3902
 8715 GGGAGACGGGGTTCACCATGTAGCCAAAGATGCTGAGATCTCTGACCTCATG--ATC 8658
 3903 CCCCCTGCTTACCTCCCAAGGGGCTGGATTAAGGTGAGTCACTGACACTTGACCAA 3962
 8657 CGCCCTCCTTGGCTCCCAAGATGCTGAGATTAAGGATGAGCCACCGCGC----- 8606
 3963 CCACATGATCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4022
 8605 CCAGCTTCTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 8546
 4023 GCGTGAAGTCAAGTGGGGCAATCTTGGCTCACTGATCACTCTGCTCCAGGTGCAAGC 4082
 8545 GCGTGAAGTCAAGT--GGCGCATCTTGGCTTACTGCAACCTCCACCTCCAGGTGCAAGC 8487
 4083 GATTTCCTGCTTACGCTCTGAGTCAAGTCAATTAAGGACACACACCAACCGCTG 4142
 8486 AATTTCCTGCTTACGCTCTGAGTCAAGTCAATTAAGGACACACCAACCGCTG 8427
 4143 CTAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 4202
 8426 CTAATGTTT-----TGTATTTTATGATTAATGATGATGATGATGATGATGATGATG 8376
 4203 TGGTCTTGAACCTCTACTCAAGTATCAACCAACCTGCGCTCCCAAGTCTGGAT 4262
 8375 TGGTCTGGAACCTCTACTCAAGTATCAACCAACCTGCGCTCCCAAGTCTGGAT 8316
 4263 TACAGTCTCAGCAGCATGACAGACGCTGATGATGATGATGATGATGATGATGATG 4322
 8315 TACAGTCTGATCAGCAGCATGACAGCTGATGATGATGATGATGATGATGATGATG 8269
 4323 TAAATGTTTATCTAAGGCACTAGCATGATGATGATGATGATGATGATGATGATG 4382
 8268 -----GAGCGCGGCGCAGTGGCTCAGCTGATGATGATGATGATGATGATGATG 8223
 4383 GGCAGAGTGGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4439
 8222 GGTCAAGATGGTGGATCAAGGATGATGATGATGATGATGATGATGATGATGATGATG 8163
 4440 CCCCCTCTACCAAAATTTAAAAATTTAGCTGGAGTGGTGGCTTGGCTGGCTG 4499
 8162 ACCCTCTCTAGTAAAA--TACAAAAATTTAGCTGGAGTGGTGGCTTGGCTGGCTG 8104
 4500 CAGTACTTGGGAAGCTAGCTGTTGGGAGTGGTGAACCTGAGTGGTGGCTGGCTG 4559
 8103 CAGTACTTGGGAAGCTAGCTGTTGGGAGTGGTGAACCTGAGTGGTGGCTGGCTG 8044
 4560 GAGTATGATCAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 4617
 8043 GAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7984
 4618 AAAATATGTTTATCTAAGGATGATGATGATGATGATGATGATGATGATGATGATG 4677
 7983 AAAATATGTTTATCTAAGGATGATGATGATGATGATGATGATGATGATGATGATG 7924
 4678 TGAATATGTTTATCTAAGGATGATGATGATGATGATGATGATGATGATGATGATG 4737

DB 7923 ATCCAGTACTCGGAGGCTGAGGAGAAATCTTTGAACTGAGAGTGAATTTG 7864
 QY 4738 C--TTAATAGCTAGACACAGACATTTGTTTCTTACTAGGACCAAGTACTTTA 4795
 DB 7863 CAGTTGCTGAGATCTACACATCTGCTTCCAGCTTGGGCAACAGAGATTCATCTC 7804
 QY 4796 AAAAATTTAGGCTCAGCAGCAGGCTGACACCTGTAATCCAGCCTTTGGAGGCA 4855
 DB 7803 AAAAATTTAGGCTCAGCAGGCTGACACCTGTAATCCAGCCTTTGGAGGCA 7744
 QY 4856 AGCAGAGAGATCATCTTGAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 4915
 DB 7743 AGCAGAGATGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 7686
 QY 4916 TCTGCTCTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4975
 DB 7685 CCCCCTCTCTACTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 7626
 QY 4976 AGCAGAGAGATCATCTTGAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 4999
 DB 7625 AGCTTTTCAAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 7602

RESULT 8
 AAS36099/c
 ID AAS36099 standard; DNA; 17792 BP.

AAS36099;
 17-DEC-2001 (first entry)

Human cardiovascular system antigen genomic DNA seq ID No 1599.

Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
 chicken; sheep; immunosuppressive; antitubercular; vasotropic; dog;
 antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
 cerebroprotective; nocitropic; antibacterial; virucide; fungicide; cancer;
 ophthalmological; vulnerrary; gene therapy; autoimmune disease; neoplasm;
 cerebrovascular disorder; breast; liver; cardiovascular disorder; ds;
 KX fungal infection; viral infection; nervous system disorder; bacterial infection;
 KX gastrointestinal disorder; renal disorder; ocular disorder; endocrine disorder;
 KX wound healing; skin aging; organ transplantation; tissue regeneration;
 KX anti-fertility.

OS Homo sapiens.

PN WO20015321-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001MO-0501340.

XX 31-JAN-2000; 2000US-0179065;
 PR 04-FEB-2000; 2000US-0180628;
 PR 24-FEB-2000; 2000US-0184664;
 PR 02-MAR-2000; 2000US-0186350;
 PR 16-MAR-2000; 2000US-0189874;
 PR 17-MAR-2000; 2000US-0190076;
 PR 18-APR-2000; 2000US-0198123;
 PR 19-MAY-2000; 2000US-0205151;
 PR 07-JUN-2000; 2000US-0209467;
 PR 28-JUN-2000; 2000US-0214886;
 PR 30-JUN-2000; 2000US-0215135;
 PR 07-JUL-2000; 2000US-0216647;
 PR 07-JUL-2000; 2000US-0218880;
 PR 11-JUL-2000; 2000US-0217487;
 PR 11-JUL-2000; 2000US-0217496;
 PR 14-JUL-2000; 2000US-0218290;
 PR 26-JUL-2000; 2000US-0220963;
 PR 26-JUL-2000; 2000US-0220964;
 PR 14-AUG-2000; 2000US-0224518;
 PR 14-AUG-2000; 2000US-0224519;

PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225457.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231142.
 PR 08-SEP-2000; 2000US-0231143.
 PR 08-SEP-2000; 2000US-0231144.
 PR 08-SEP-2000; 2000US-0231145.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0234998.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241185.
 PR 20-OCT-2000; 2000US-0241186.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249246.
 PR 17-NOV-2000; 2000US-0249255.
 PR 17-NOV-2000; 2000US-0249257.
 PR 17-NOV-2000; 2000US-0249259.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251899.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM.
 MPI; 2001-451930/48.
 New cardiovascular system related polynucleotides and polypeptides,
 useful for diagnosing, treating and/or preventing disorders of the
 cardiovascular system -
 Claim 1, SEQ ID No 1599; 674pp; English.
 Sequences AAS35741-AAS36942 represent genomic DNA molecules, which encode
 the cardiovascular system antigen polypeptides of the invention.
 Cardiovascular system antigens and their associated polynucleotides are
 useful in the diagnosis, treatment and prevention of various types of
 disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
 chickens or sheep. A pathological condition can be determined by
 detecting the presence or absence of a mutation in a cardiovascular
 system antigen polynucleotide. The treatable disorders include autoimmune
 diseases such as rheumatoid arthritis, hyperproliferative disorders such
 as neoplasms of the breast or liver, cardiovascular disorders such as
 cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,
 nervous system disorders such as Alzheimer's disease, infections caused
 by bacteria, viruses and fungi, ocular disorders such as corneal
 infection, endocrine disorders such as premature labour and infertility,
 gastrointestinal disorders such as Crohn's disease, renal disorders such
 as glomerulonephritis and respiratory disorders such as asthma and
 pleurisy. The polypeptides can also be used to aid wound healing, to


```
Db 1708 ACCTGCTCAGCCTCCAAAGTCTAGATTAACAGGCTGAG---CACACCGGCGCG 1764
Qy 3964 CACAGGATCTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACAGGCTTCACTCCATCACCAG 4023
Db 1765 CCGTCACAACTTTTTTTTTTTTTTTTTTTTTTTTGTGAGAGAGTCTTGCTGCTGACTGCG 1824
Qy 4024 GCTGAGTACAGTGGGGGCAATCTGGCTCAGTAACTTAACTTGGCTCCAGGTCGCAACG 4083
Db 1825 GCTGAGGAGAGT-GGCCCCGATCTCAGTCTCAGTCACTCAACCTCCGCTCTGCTGATGCG 1883
Qy 4084 ATTCTCTGCTTAACTCTCTGAGTACGATGGAATTAAGCAGACACACAGCCGCGG 4143
Db 1884 CTCTCTGCTCTCAATCTCCCAAGTACGAGTATAGGCGCCACACACGCGCGG 1943
Qy 4144 TAATTTTTTTTTTTTTTTCTGATTTTCTAGTATAGACAGGCTTCACTATGTCGCAAGCT 4203
Db 1944 TAATTTTT-----TGTATTTTATGAGAGAGGCGGTTTCACTGTTTACGAGAT 1994
Qy 4204 GGTCTTAACCCCTGACCTCAAGTATCAACCCAGCTCGGCTCCCAAGTGGGAT 4263
Db 1995 GGTCTCATGCGCTGACCTC--GTATCGCGCGGCTCACCTTACCAAGTGGGAT 2052
Qy 4264 ACAGGTCTCAGCCACCATGACAGCCACATGCTACTTTTAAATTTATTTTAA 4323
Db 2053 ACAGGTCTGAGCCACCTGCGCGCCCAACTATTTTAAACCTTATGCTTAA 2112
Qy 4324 AAAATGTTATCTAAGCCAGTACAGTACGCTGCTGTAATCCAGACTTTAGG 4383
Db 2113 ACTAGAGCATGACAGAGCTGGGCGTGGCTGACCTCTTATCCAGACTTTGGAG 2172
Qy 4384 GCCAGGTGGGGGATCAGTGGCTGGAGT-----CAGCGTGGCAATAGTGA 4437
Db 2173 GCGAGTGGGTGATACCTGAGGTGGAGTTGAGACAGCTGAGCAACATGAGGA 2232
Qy 4438 GACCCGCTCTACCAAAATTTAAAAATTAGCTGGAGTGGGCAATTTGCTGT 4497
Db 2233 AACCCGCTCTACTGAAAA--TACAAATTAGCCGGGTGTGTGGCAATGCTGTAT 2290
Qy 4498 CCCAGTACTTGGGAAGCTGAGGTGTGGGATGCTGAAGCCTGAGGTGAGGCTGCA 4557
Db 2291 CCCAGTACTTGGAGGCTGAGGAGGAGGAATCGCTTGAACCAAGAGGAGAAATTCG 2350
Qy 4558 GTGAGCTATGATCACACACTGCACTTCACTGCTGATACAG 4600
Db 2351 GTGAGCCGAGATCGGCGCATTCGACTCAAGCTGCGCAACAG 2393

RESULT 10
AKK84002/c
ID AKK84002 standard; DNA; 15413 BP.
XX
AC AKK84002;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38814.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX
KW Cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PM MO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR
```

```
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216547.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR
```


DB 42419 CCAGCTACTTGGGAGGCTGAGAGGAGATCACTTGAATCTGGAGGAGAGGTTGCAG 42360
OY 4559 TGAGCTGATGATCACACCATGCTGACCTGAGTGCAC 4597
DB 42339 TGAGCTGAGTTACACCATGCTGACCTGAGTGCAC 42321

RESULT 12

AAK89986
ID AAK89986 standard; DNA: 32224 BP.

AC AAK89986;

DT 05-NOV-2001 (first entry)

DE Human digestive system antigen genomic sequence SEQ ID NO: 3562.

KM Human: digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
XX digestive system disorder; Meckel's diverticulum; ds.

OS Homo sapiens.

PN WO200155314-A2.

PD 02-AUG-2001.

PE 17-JAN-2001; 2001WO-US01324.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0180628.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217487.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 14-AUG-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 18-AUG-2000; 2000US-0225759.

PR 23-AUG-2000; 2000US-0226279.

PR 23-AUG-2000; 2000US-0226681.

PR 23-AUG-2000; 2000US-0226681.

PR 23-AUG-2000; 2000US-0227182.

PR 30-AUG-2000; 2000US-0227009.

PR 01-SEP-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 05-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 06-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0233397.

PR 14-SEP-2000; 2000US-0233399.

PR 14-SEP-2000; 2000US-0233399.

PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 21-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.

PR 29-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236337.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 13-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241121.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.

PR 01-NOV-2000; 2000US-0241826.

PR 08-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246533.

PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.

PR 17-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.

PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.

PR 17-NOV-2000; 2000US-0249212.

PR 17-NOV-2000; 2000US-0249213.

PR 17-NOV-2000; 2000US-0249214.

PR 17-NOV-2000; 2000US-0249215.

PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.

PR 17-NOV-2000; 2000US-0249218.

The present invention provides the protein and coding sequences of a number of human digestive system antigens. These can be used in the diagnosis, treatment and prevention of digestive system disorders, including cancer, Crohn's disease, celiac disease, bacterial or parasitic infections, appendicitis, Hirschsprung's disease, chronic colitis, ulcerative colitis. The present sequence is a genomic DNA fragment encoding a digestive system antigen of the invention.

```

          3407: Conservative      0: Mismatches 173: Indels      41: Gaps      5:
3844  AGAGATGAGAGTGCATCTGCTGTGTGCCAGGCTGCTCGAATCTCAAGGCTCAACCAATCC 3903
      |||||      |||||
16163  AAGAGTGGGATTTACCATATTTGCAAGCTGTGTTCCAACTCCGACATCAAGGTGTGC 16222
      |||||      |||||
3904  CCTGCTTACCTCCCAAGGGCTGGGATTACAGGTGTGACTGACTGATTGACCAAC 3963
      |||||      |||||
16223  GCCTGCTCAAGCTCCCAAGTCTGTACTACAGGATGAGCCACTGCGCTGCGC---- 16278
      |||||      |||||
3964  CACATGTACTTTTTTTTTTTTTTTTTTTTTTTTGGAGAGGGTTTCACTCATCAACCCAG 4023
      |||||      |||||
16279  --CGGTGTTGAGATTTTGTGTTTATTATTGATGAGATGAGTCCGCTGTCGCGCTAG 16336
      |||||      |||||
4024  GCTGATGAGAGGGGGCAATCTTGGCTCACTGTAACCTGTGCTCCAGGTGCAAGG 4083
      |||||      |||||
16337  GTTGAGATGAGT--GGCATGATCTCGGCTCACTCAAGCTCTCTTCTCGGGTTCAACGA 16395
      |||||      |||||
4084  ATTCTCTGAGCTTACGCTCGTAGAGTGTGAATTAATAGCACACACACAGGCTGAGC 4143
      |||||      |||||
16396  ATTCTCTGCTCAAGCTCCGAGAGTGGGACTACAGGCTCCCACTACCAACACTGTGC 16455
      |||||      |||||
4144  TAATTTTTTTTTTTTTTCTGATTTTGTAGTAGACAGGGTTTCATCATGTTGGCAGGCT 4203
      |||||      |||||
16456  TAATTTT-----TGAATTTTGTGCAACGGGGTTTTCGCAATATTTGGCAAGCT 16506
      |||||      |||||
4204  GGATCTGAACCCCTGACTCAAGTATCAACCCAGCTGGGCTCCCAAGTCTGGGATT 4263
      |||||      |||||

```

RESULT 13
AAL36231/C
ID AAL36231 standard; DNA: 13819 BP.
XX

| | | | |
|----|--|--|--|
| AC | | | |
| XX | AAI36231; | | |
| XX | | | |
| DT | 08-JAN-2002 (first entry) | | |
| XX | | | |
| XX | Human musculoskeletal system related polynucleotide SEQ ID NO 2596. | | |
| XX | | | |
| XX | Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor; vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic; cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder; neurological disease; infection; human; secreted protein; musculoskeletal system; ds. | | |

| | |
|------------------------------|---------------|
| | Homo sapiens. |
| MO200I55367-A1. | |
| 02-AUG-2001. | |
| 17-JAN-2001; 2001WO-US01338. | |
| 31-JAN-2000; 2000US-0179065. | |
| 04-FEB-2000; 2000US-0180628. | |
| 24-FEB-2000; 2000US-0184664. | |
| 02-MAR-2000; 2000US-0186350. | |
| 16-MAR-2000; 2000US-0186874. | |
| 17-MAR-2000; 2000US-0190076. | |
| 18-APR-2000; 2000US-0198123. | |
| 19-MAY-2000; 2000US-0205515. | |
| 07-JUN-2000; 2000US-0209467. | |
| 28-JUN-2000; 2000US-0214886. | |
| 30-JUN-2000; 2000US-0215135. | |
| 07-JUL-2000; 2000US-0216647. | |
| 07-JUL-2000; 2000US-0216880. | |
| 11-JUL-2000; 2000US-0217487. | |
| 11-JUL-2000; 2000US-0217496. | |
| 14-JUL-2000; 2000US-0218290. | |
| 26-JUL-2000; 2000US-0220963. | |
| 26-JUL-2000; 2000US-0220964. | |
| 14-AUG-2000; 2000US-0224518. | |
| 14-AUG-2000; 2000US-0224519. | |
| 14-AUG-2000; 2000US-0225213. | |

Mon Jun 9 10:49:15 2003

us-09-622-964-1_copy_1_5000.rng

Page 32

[illegible]

Search completed: June 8, 2003, 11:23:43
Job time : 1023.38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:

June 8, 2003, 09:30:48 ; Search time 12606.1 Seconds
(without alignments)
11543.117 Million cell updates/sec

Title: US-09-622-964-1-COPY_1_5000
Perfect score: 5000
Sequence: 1 ccaaaaatgtctctctgtg.....tttgagagccgagcggt 5000

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 4109280
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

GenEmbl: *
1: gb_da: *
2: gb_hlg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_ov: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_hlg_hum: *
31: em_hlg_in: *
32: em_hlg_inv: *
33: em_hlg_other: *
34: em_hlg_mus: *
35: em_hlg_pln: *
36: em_hlg_rtd: *
37: em_hlg_mam: *
38: em_hlg_vrt: *
39: em_hlg_hum: *
40: em_hlg_mus: *
41: em_hlg_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|--------------|----------------------|
| 1 | 4034.2 | 80.7 | 112309 | 9 AC003025 | AC003025 Human Chr |
| 2 | 4034.2 | 80.7 | 142092 | 9 AF139813 | AF139813 Homo sapi |
| 3 | 4034.2 | 80.7 | 196080 | 9 AC004228 | AC004228 Homo sapi |
| 4 | 3254.6 | 65.1 | 163915 | 2 AC087451 | AC087451 Homo sapi |
| 5 | 3254.6 | 65.1 | 166867 | 9 AP003733 | AP003733 Homo sapi |
| 6 | 3142 | 62.8 | 133683 | 2 AC084857 | AC084857 Homo sapi |
| 7 | 2379.8 | 47.6 | 160169 | 2 AC051664 | AC051664 Homo sapi |
| 8 | 1291.8 | 25.8 | 6396 | 6 AX347149 | AX347149 Sequence |
| 9 | 1284.2 | 25.7 | 6396 | 6 AX347148 | AX347148 Sequence |
| 10 | 742.2 | 14.8 | 160169 | 2 AC051664 | AC051664 Homo sapi |
| 11 | 728.2 | 14.6 | 133683 | 2 AC084857 | AC084857 Homo sapi |
| 12 | 435.6 | 8.7 | 165423 | 9 CDS07ECV | AL442663 Homo chr |
| 13 | 431 | 8.6 | 125508 | 9 AC006345 | AC006345 Homo sapi |
| 14 | 406.2 | 8.1 | 98589 | 9 HSD309K20 | AL109827 Homo sapi |
| 15 | 405.4 | 8.1 | 121502 | 9 AC007136 | AL050307 Homo sapi |
| 16 | 398.6 | 8.0 | 42265 | 9 AC007136 | AL050307 Homo sapi |
| 17 | 394.8 | 7.9 | 142000 | 9 AP000845 | AP000845 Homo sapi |
| 18 | 394.6 | 7.9 | 186335 | 2 AC068889 | AC068889 Homo sapi |
| 19 | 392 | 7.8 | 166869 | 9 AC016395 | AL163579 Homo sapi |
| 20 | 387.6 | 7.7 | 134010 | 9 AL353579 | AL353579 Homo sapi |
| 21 | 387.2 | 7.7 | 100827 | 9 AC020909 | AC020909 Homo sapi |
| 22 | 385.2 | 7.7 | 70693 | 9 AL592314 | AL592314 Homo sapi |
| 23 | 384.8 | 7.7 | 11046 | 9 AC124612 | AC124612 Homo sapi |
| 24 | 384.6 | 7.7 | 11848 | 2 AC090574 | AC090574 Homo sapi |
| 25 | 384.6 | 7.7 | 172969 | 9 AC092066 | AC092066 Homo sapi |
| 26 | 383 | 7.7 | 137411 | 9 AL355490 | AL355490 Homo sapi |
| 27 | 382.2 | 7.6 | 190803 | 9 AC112717 | AC112717 Homo sapi |
| 28 | 382 | 7.6 | 189406 | 9 AC055726_2 | AC055726_2 Homo sapi |
| 29 | 380.6 | 7.6 | 110000 | 2 AC023050 | AC023050 Homo sapi |
| 30 | 380.6 | 7.6 | 191474 | 2 AC023157 | AC023157 Homo sapi |
| 31 | 380.6 | 7.6 | 210544 | 2 AC013359 | AC013359 Homo sapi |
| 32 | 380.2 | 7.6 | 171812 | 9 AC079988 | AC079988 Homo sapi |
| 33 | 380.2 | 7.6 | 211305 | 9 AC005102 | AC005102 Homo sapi |
| 34 | 380 | 7.6 | 106508 | 9 AC090063 | AL606465 Homo sapi |
| 35 | 378.8 | 7.6 | 173391 | 9 AL583963 | AL583963 Homo sapi |
| 36 | 378.4 | 7.6 | 116103 | 9 AL583963 | AL583963 Homo sapi |
| 37 | 378.2 | 7.6 | 80146 | 9 HS1112F19 | AL158040 Homo sapi |
| 38 | 378.2 | 7.6 | 116368 | 9 HS1112F19 | AL158040 Homo sapi |
| 39 | 378.2 | 7.6 | 195370 | 2 AL158040 | AL158040 Homo sapi |
| 40 | 378.2 | 7.6 | 213648 | 9 HS215D11 | AL158040 Homo sapi |
| 41 | 377 | 7.5 | 122379 | 2 AL512402 | AL512402 Homo sapi |
| 42 | 377 | 7.5 | 223597 | 2 AC004517 | AC004517 Homo sapi |
| 43 | 376.8 | 7.5 | 44062 | 9 AC004686 | AC004686 Homo sapi |
| 44 | 376.6 | 7.5 | 133478 | 9 AC097359 | AC097359 Homo sapi |
| 45 | 375.6 | 7.5 | 187127 | 9 AC097359 | AC097359 Homo sapi |

ALIGNMENTS

| RESULT 1 | AC003025 | 112309 bp | DNA | linear | PRI 23-JUL-1998 |
|------------|--|------------|-----|--------|-----------------|
| LOCUS | AC003025 | | | | |
| DEFINITION | Human Chromosome 11p12.2 PAC clone pDU466a11, complete sequence. | | | | |
| ACCESSION | AC003025 | | | | |
| VERSION | AC003025.1 | GI:3337308 | | | |
| KEYWORDS | HTG. | | | | |
| SOURCE | Homo sapiens | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | |
| AUTHORS | Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basti,M., Buettner,J., Bunnister,R., Card,P., Desalpoat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., | | | | |

TITLE
JOURNAL
REFERENCE
AUTHORS
Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, H., Loo, K.N., Major, T., McFarland, J., Newton, J., Osborne-Lawrence, S., Schageman, J., Schultz, R.A., Stimson, S., Syed, M. and Ward, T.
HGS Submission
Unpublished
2 (bases 1 to 112309)

TITLE
JOURNAL
REFERENCE
AUTHORS
Submitted (21-Oct-1997) Genome Science and Technology Center, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
3 (bases 1 to 112309)
Evans, G.A., Athanasiou, M., Bradbury, P., Brianac, S., Bumeister, R., Davis, C., English, C., Franklin, T.L., Garner, H.R., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J., Hinson, S., Naryanaswamy, U., Newton, J., O'Brien, K., Oliver, T., Patel, P., Probst, S., Rayner, S., Schageman, J., Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T. and Wilson, R.
Direct Submission

TITLE
JOURNAL
COMMENT
Submitted (23-Jun-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
On Jun 23, 1998 this sequence version replaced g1:2554967.
IMPORTANT: This submission contains the entire insert of clone PDJ466a11. PDJ466a11 comes from a PAC library constructed at the Roswell Park Cancer Institute by the Pieter de Jong group. This clone has been finished according to strict quality criteria and attempts have been made to resolve all base calling problems such as compressions and repetitive elements. The expected Phred/Phrap calculated errors/10kb is 0.18. In addition, this sequence has been finished such that 99.9% of consensus base calls consist of either double-stranded coverage or 2 types of labeling chemistry on one strand.
Further information regarding the map of this region or annotation of PDJ466a11 can be found at
http://gsstec.swmed.edu/chromosome1.htm.
CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11p12.2 Best's disease region mapped between STS D11S461 and EST A1NAK. This region spans over 1.5 Mbp.
MARKER CONFIRMATION: EST: FTH (ferritin heavy chain mRNA)
MAPPED CLONE OVERLAP: HGS submitted PAC clones PDJ519013 and PDJ756b9.

FEATURES
SOURCE
Location/Qualifiers
1..112309

repeat_region /organism="Homo sapiens"
repeat_region /db_xref="taxon:9606"
repeat_region complement(4341..4656)
repeat_region /rpt_family="Alu"
repeat_region complement(9880..10120)
repeat_region /rpt_family="Alu"
repeat_region 10163..10282
repeat_region /rpt_family="Alu"
repeat_region complement(11232..12157)
repeat_region /rpt_family="Alu"
repeat_region 13351..13560
repeat_region /rpt_family="L1"
repeat_region complement(15092..15146)
repeat_region /rpt_family="MIR"
repeat_region 15826..15932
repeat_region /rpt_family="MIR"
repeat_region 17011..17146
repeat_region /rpt_family="MIR"
repeat_region complement(18018..18313)
repeat_region /rpt_family="Alu"
repeat_region 18493..18741
repeat_region /rpt_family="Alu"
repeat_region 18820..18954
repeat_region /rpt_family="Alu"

repeat_region complement(21374..21661)
repeat_region /rpt_family="Alu"
repeat_region 22809..23212
repeat_region /rpt_family="Alu"
repeat_region 23499..23752
repeat_region /rpt_family="Alu"
repeat_region 24118..24472
repeat_region /rpt_family="Alu"
repeat_region 25005..25113
repeat_region /rpt_family="MIR"
repeat_region 31160..31442
repeat_region /rpt_family="Alu"
repeat_region complement(32651..32947)
repeat_region /rpt_family="Alu"
repeat_region complement(35970..36025)
repeat_region /rpt_family="MIR"
repeat_region 38958..39145
repeat_region /rpt_family="Alu"
repeat_region 44878..45480
repeat_region /rpt_family="Alu"
repeat_region 49769..50071
repeat_region /rpt_family="Alu"
repeat_region complement(50172..50310)
repeat_region /rpt_family="Alu"
repeat_region 50902..50982
repeat_region /rpt_family="MIR"
repeat_region 51038..51659
repeat_region /rpt_family="Alu"
repeat_region 54494..54769
repeat_region /rpt_family="Alu"
repeat_region 55529..55623
repeat_region /rpt_family="THE1"
repeat_region complement(55691..55967)
repeat_region /rpt_family="Alu"
repeat_region complement(56049..56257)
repeat_region /rpt_family="Alu"
repeat_region 56721..57100
repeat_region /rpt_family="THE1"
repeat_region 58402..58680
repeat_region /rpt_family="Alu"
repeat_region complement(58748..59345)
repeat_region /rpt_family="Alu"
repeat_region 59678..59904
repeat_region /rpt_family="Alu"
repeat_region 61149..61424
repeat_region /rpt_family="Alu"
repeat_region complement(61930..62242)
repeat_region /rpt_family="Alu"
repeat_region 62276..62710
repeat_region /rpt_family="MLT1"
repeat_region 62915..63118
repeat_region /rpt_family="Alu"
repeat_region complement(63349..63455)
repeat_region /rpt_family="Alu"
repeat_region 64339..64614
repeat_region /rpt_family="Alu"
repeat_region complement(64707..64984)
repeat_region /rpt_family="Alu"
repeat_region complement(65596..65811)
repeat_region /rpt_family="Alu"
repeat_region complement(66500..66753)
repeat_region /rpt_family="Alu"
repeat_region 67311..67342
repeat_region /rpt_family="Alu"
repeat_region 67835..68119
repeat_region /rpt_family="Alu"
repeat_region 68860..69150
repeat_region /rpt_family="Alu"
repeat_region 69164..69272
repeat_region /rpt_family="Alu"
repeat_region 69354..70245
repeat_region /rpt_family="Alu"
repeat_region 70655..70939

```
repeat_region /rpl_family="Alu" 70933..71075
repeat_region /rpl_family="L1" 71119..71395
repeat_region /rpl_family="Alu" 71404..71530
repeat_region /rpl_family="LTR5" 71536..71637
repeat_region /rpl_family="LTR5" 72720..72948
repeat_region /rpl_family="L1" complement(73488..73633)
repeat_region /rpl_family="Alu" complement(73925..74200)
repeat_region /rpl_family="Alu" 75721..75955
repeat_region /rpl_family="Alu" 77185..77449
repeat_region /rpl_family="Alu" 78093..78222
repeat_region /rpl_family="L1" 78227..78294
repeat_region /rpl_family="L1" complement(80458..80707)
repeat_region /rpl_family="Alu" complement(81544..81974)
repeat_region /rpl_family="Alu" 82055..82305
repeat_region /rpl_family="Alu" 82514..83096
repeat_region /rpl_family="Alu" complement(83685..83940)
repeat_region /rpl_family="Alu" 85536..85850
repeat_region /rpl_family="Alu" complement(86429..86483)
repeat_region /rpl_family="MIR" 86902..87183
repeat_region /rpl_family="Alu" 8746..87792
repeat_region /rpl_family="Alu"

Query Match 80.7% Score 4034.2 DB 9 Length 112309;
Best Local Similarity 91.9% Pred. No. 0;
Matches 4486; Conservative 0; Mismatches 28; Indels 365; Gaps 7;

QY 122 GCAACCCACATGCGACAGCTATCTATGTAAACCACTGCACATCATGTATACCTAT 181
DB 78181 GCAACCCACATGCGACAGCTATCTATGTAAACCACTGCACATCATGTATACCTAT 181
QY 182 GTAACCACTGGTATCTTGTGACAGCTATCCAGACTTTAGAGTGAAGAAAAAGTG 241
DB 78240 GTAACCACTGGTATCTTGTGACAGCTATCCAGACTTTAGAGTGAAGAAAAAGTG 241
QY 242 GTGTGTAGAAAAATTCCTGCAATCTCAGCATAGTTAAGCTTATGATACCTTTGAGAGA 301
DB 78298 GTGTGTAGAAAAATTCCTGCAATCTCAGCATAGTTAAGCTTATGATACCTTTGAGAGA 301
QY 302 GAGGGTGACAGAAAGGAGATGAGAGTGGTTAAAGACAAAGTCAATTTAAAT 361
DB 78358 GAGGGTGACAGAAAGGAGATGAGAGTGGTTAAAGACAAAGTCAATTTAAAT 361
QY 362 CAGGGCTCTGGAAGTTATGTCACAAAACACATCTCATATATCCCTCAGTGTGA 78417
DB 78418 CAGGGCTCTGGAAGTTATGTCACAAAACACATCTCATATATCCCTCAGTGTGA 78417
QY 422 TTAATATGCAACATCCCTAAGGCCACAGACTCTGAGAAAGATCCAGAAAAAGT 481
DB 78478 TTAATATGCAACATCCCTAAGGCCACAGACTCTGAGAAAGATCCAGAAAAAGT 481
QY 482 CCCGTTAATTAACATTTGGCGATTTCTTACGGCTCTTAAGACCAAGAACACATGCTGC 541
DB 78538 CCCGTTAATTAACATTTGGCGATTTCTTACGGCTCTTAAGACCAAGAACACATGCTGC 541
```

```
QY 542 CTAGAGCTCTGCTCTCTTCAATTAACATATACAGAGAGTGTAGGTAGACACCCACCA 601
DB 78598 CTAGAGCTCTGCTCTCTTCAATTAACATATACAGAGAGTGTAGGTAGACACCCACCA 601
QY 602 CTTCACACAGCTTAGAGAGAGCCCTTGATATGATATATTAATTAATTAATGAATCA 661
DB 78658 CTTCACACAGCTTAGAGAGAGCCCTTGATATGATATATTAATTAATTAATGAATCA 661
QY 662 CATCGTAGATTTTACACAGCTCTCCGCGGAGATCTGGGCATTTATCCCATTTGCACT 721
DB 78718 CATCGTAGATTTTACACAGCTCTCCGCGGAGATCTGGGCATTTATCCCATTTGCACT 721
QY 722 GCGTGGCTGGAAAGCCAGACATTAACATCCAGAGCTGTCTGTCAACCCACAGACTC 78777
DB 78778 GCGTGGCTGGAAAGCCAGACATTAACATCCAGAGCTGTCTGTCAACCCACAGACTC 78777
QY 782 ACCCCCTCCACAGCCCGGCGAGCTTCTCTTCATCTCTGTAAGCACTTATGAT 841
DB 78838 ACCCCCTCCACAGCCCGGCGAGCTTCTCTTCATCTCTGTAAGCACTTATGAT 841
QY 842 GGGCCCTCCAGCCCATACAGCCAGAAATTAACATATGATGTACACAGACCAATACAG 901
DB 78898 GGGCCCTCCAGCCCATACAGCCAGAAATTAACATATGATGTACACAGACCAATACAG 901
QY 902 CTCCGTCAGCATATGACAGAAATTTCTCATTTTACTAGGGTGTGAAATTTCCAAAGCA 961
DB 78958 CTCCGTCAGCATATGACAGAAATTTCTCATTTTACTAGGGTGTGAAATTTCCAAAGCA 961
QY 962 CACCATCTTTTACAGATTAAGGAGCTGAGCTGAGAGAGAGTGAACCTACCCGGGT 1021
DB 79018 CACCATCTTTTACAGATTAAGGAGCTGAGCTGAGAGAGAGTGAACCTACCCGGGT 1021
QY 1022 CACACACACAGCTGAGAGAGCTGGGACCAAGAAACAGAGCTGTGACAGCCCGGTA 1081
DB 79078 CACACACACAGCTGAGAGAGCTGGGACCAAGAAACAGAGCTGTGACAGCCCGGTA 1081
QY 1082 TTCTATCTTTTCAATGAGCCACAGAGCTGTGCAAAAGCCAGAGCTTACAGAGCTCT 1141
DB 79138 TTCTATCTTTTCAATGAGCCACAGAGCTGTGCAAAAGCCAGAGCTTACAGAGCTCT 1141
QY 1142 CCTTCTGAGAGCTTCTGCAACAGAAATTAACAGCTTACAGCCCTTAAACCCCAAC 1201
DB 79198 CCTTCTGAGAGCTTCTGCAACAGAAATTAACAGCTTACAGCCCTTAAACCCCAAC 1201
QY 1202 TCTCTGCAAGAGCTCAGAGGAGTGAAGAACTGTGAGAGAGATCTTTAGCTTGGAT 1261
DB 79258 TCTCTGCAAGAGCTCAGAGGAGTGAAGAACTGTGAGAGAGATCTTTAGCTTGGAT 1261
QY 1262 TTATAGGCGCATGTGAGAGGAGTGTGCTTAAATTCAGAGCCCTGCTCAAGCAAC 1321
DB 79318 TTATAGGCGCATGTGAGAGGAGTGTGCTTAAATTCAGAGCCCTGCTCAAGCAAC 1321
QY 1322 CTTCCAAAGAAATTAAGAGGAGCTGAGCCAGAGCTGTCTAGCCGTTCTCTGAGAG 1381
DB 79378 CTTCCAAAGAAATTAAGAGGAGCTGAGCCAGAGCTGTCTAGCCGTTCTCTGAGAG 1381
QY 1382 ATTACAAAGAGAGCTTAAGACAGATCTCTTGTGAGAGTCTGAGAGTCAAG 1441
DB 79438 ATTACAAAGAGAGCTTAAGACAGATCTCTTGTGAGAGTCTGAGAGTCAAG 1441
QY 1442 TGAAGCGGCTCAGACTCAGTGGGAGTCCAGACCTCTAAGAGTGGAGGGGCACTG 1501
DB 79498 TGAAGCGGCTCAGACTCAGTGGGAGTCCAGACCTCTAAGAGTGGAGGGGCACTG 1501
QY 1502 GCCACAGAGTCCAGAGAGTCCACACAGCTTATGCGCAGACCTCTGTTGGATCATGCG 1561
DB 79558 GCCACAGAGTCCAGAGAGTCCACACAGCTTATGCGCAGACCTCTGTTGGATCATGCG 1561
QY 1562 ACCCAGCTGAACCCCACTGTGAGTCAAGAGTCCCAAGTGTGAGTGGGCTGCTT 1621
DB 79618 ACCCAGCTGAACCCCACTGTGAGTCAAGAGTCCCAAGTGTGAGTGGGCTGCTT 1621
```

[illegible][illegible]

81476 CCCACACGACTCATAGGCCACATAGTACATTAAAAAGAGAGAGAGAGAGAGAG 81535

8842 AGAGATGGAGTCTACTGTGTGTGTGCAGGCGTGTGGAACTCTAGCTTCAGCAAT 81535

81536 AGAGATGGAGTCTACTGTGTGTGTGCAGGCGTGTGGAACTCTAGCTTCAGCAAT 3901

3902 CCCCGTCCCTTACGCTCCCAAGGGCTGGAGTTACAGGTGTGAGTACTGCACTTGACCA 81595

81536 CCCCGTCCCTTACGCTCCCAAGGGCTGGAGTTACAGGTGTGAGTACTGCACTTGACCA 3961

3962 ACCACATGGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGAGACAGGGTTCTACTCATCACC 81655

81656 ACCACATGGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGAGACAGGGTTCTACTCATCACC 4021

4022 AGCGTGAAGTCAAGTGGGGCAATCTTGCTCACTGTAACTCTGCGCTCCAGAGTGAAG 81715

81716 AGCGTGAAGTCAAGTGGGGCAATCTTGCTCACTGTAACTCTGCGCTCCAGAGTGAAG 4081

4082 CGATTCCTCCGCTTACGCTCCCAAGGGCTGGAGTTACAGGTGTGAGTACTGCACTTGACCA 81775

81776 CGATTCCTCCGCTTACGCTCCCAAGGGCTGGAGTTACAGGTGTGAGTACTGCACTTGACCA 4141

4142 GCTAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGATTTTGTAGAGACAGGGTTTCATCATGTGGCCAG 81835

81836 GCTAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGATTTTGTAGAGACAGGGTTTCATCATGTGGCCAG 4201

4202 CAGGCTTCGACCCCTGACCTCACTCACTGATGATCCACCCACCTCGGCTCCCAAGTGTGGGA 81895

81896 CAGGCTTCGACCCCTGACCTCACTCACTGATGATCCACCCACCTCGGCTCCCAAGTGTGGGA 4261

4262 TTACAGGTGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 81955

81956 TTACAGGTGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4321

4322 TTAAGT 82015

82016 TTAAGT 4381

4382 GGGCCAGGAGGCGGGGATCACTGTAGCCTGGAGTTACAGGTGTGAGTGTGAGTGTGAGC 82075

82076 GGGCCAGGAGGCGGGGATCACTGTAGCCTGGAGTTACAGGTGTGAGTGTGAGTGTGAGC 4441

4442 CGGTCTTACAAAAATTTTAAAAATTTTAACTGGAGTGTGAGTGTGAGTGTGAGTGTGAGC 82135

82136 CGGTCTTACAAAAATTTTAAAAATTTTAACTGGAGTGTGAGTGTGAGTGTGAGTGTGAGC 4501

4502 GCTACTTGGGAAAGCTGAGGCTGTGGGATGGCTGAAAGCTGTGAGTGTGAGTGTGAGTGTGAGC 82195

82196 GCTACTTGGGAAAGCTGAGGCTGTGGGATGGCTGAAAGCTGTGAGTGTGAGTGTGAGTGTGAGC 4561

4562 GCTACTTGGGAAAGCTGAGGCTGTGGGATGGCTGAAAGCTGTGAGTGTGAGTGTGAGTGTGAGC 82255

82256 GCTACTTGGGAAAGCTGAGGCTGTGGGATGGCTGAAAGCTGTGAGTGTGAGTGTGAGTGTGAGC 4621

4622 TAAATGTATCTAAACGATGAGTATATACAGGAATATGATATGATATGATATGATATGATATGAT 82315

82316 TAAATGTATCTAAACGATGAGTATATACAGGAATATGATATGATATGATATGATATGATATGAT 4681

4682 TAAATGTATCTAAACGATGAGTATATACAGGAATATGATATGATATGATATGATATGATATGAT 82375

82376 TAAATGTATCTAAACGATGAGTATATACAGGAATATGATATGATATGATATGATATGATATGAT 4741

4742 TAAATGTATCTAAACGATGAGTATATACAGGAATATGATATGATATGATATGATATGATATGAT 82435

82436 TAAATGTATCTAAACGATGAGTATATACAGGAATATGATATGATATGATATGATATGATATGAT 4801

4802 GTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 82495

82496 GTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4861

4862 GAGATCACTTGAAGCCAGAGGATTTAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 82555

82556 GAGATCACTTGAAGCCAGAGGATTTAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 82615


```
repeat_region /rpt_family="SVA" complement(5503, .6111)
repeat_region /rpt_family="Alu" 5569, .6095
repeat_region /rpt_family="SVA" complement(7239, .7542)
repeat_region /rpt_family="Alu" 7306, .7497
repeat_region /rpt_family="SVA" 7732, .7828
repeat_region /rpt_family="MIR" 9099, .9403
repeat_region /rpt_family="Alu" complement(9526, .9897)
repeat_region /rpt_family="MER1" 9635, .9752
repeat_region /rpt_family="MER1B" 11924, .12197
repeat_region /rpt_family="Alu" complement(11955, .12151)
repeat_region /rpt_family="SVA" complement(12767, .13466)
repeat_region /rpt_family="Alu" 16757, .17039
repeat_region /rpt_family="Alu" complement(16775, .17010)
repeat_region /rpt_family="SVA" complement(18220, .18507)
repeat_region /rpt_family="Alu" 18329, .18485
repeat_region /rpt_family="SVA" 20101, .23446
gene /note="HUMFERRITIN" join(20101, .20421, 22218, .22364, 22621, .22746, 22842, .23446)
mRNA /gene="HUMFERRITIN" /product="ferritin heavy chain"
repeat_region /rpt_family="Alu" 21506, .21714
repeat_region /rpt_family="SVA" complement(23272, .37375)
gene /gene="VMD2" /note="Best macular dystrophy gene" complement(join(23372, .23629, 24840, .25478, 27690, .27841, 28155, .28235, 29427, .29579, 30274, .30351, 30740, .30894, 31787, .32030, 32537, .32631, 35777, .35964, 37308, .37375))
/ gene="VMD2" /product="Best macular dystrophy protein" complement(23735, .23954)
repeat_region /rpt_family="Alu" 23789, .23956
repeat_region /rpt_family="SVA" 26429, .26707
repeat_region /rpt_family="Alu" complement(26444, .26696)
repeat_region /rpt_family="SVA" 27103, .27349
repeat_region /rpt_family="Alu" complement(27106, .27323)
repeat_region /rpt_family="SVA" complement(29161, .29307)
repeat_region /rpt_family="Alu" complement(29762, .30043)
repeat_region /rpt_family="Alu" 29786, .30017
repeat_region /rpt_family="SVA" 30462, .30516
repeat_region /rpt_family="MIR" complement(31095, .31409)
repeat_region /rpt_family="Alu" 33005, .33260
repeat_region /rpt_family="Alu"
```

```
repeat_region 33247, .33597 /rpt_family="L1MC2"
repeat_region complement(33849, .34431) /rpt_family="Alu"
repeat_region 34076, .34411 /rpt_family="SVA"
repeat_region complement(34640, .34890) /rpt_family="Alu"
repeat_region 34972, .35401 /rpt_family="Alu"
repeat_region complement(34988, .35392) /rpt_family="SVA"
repeat_region 36238, .36487 /rpt_family="Alu"
repeat_region complement(36253, .36457) /rpt_family="SVA"
repeat_region complement(38661, .38855) /rpt_family="L1"
repeat_region complement(40397, .41224) /rpt_family="Alu"
repeat_region 41056, .41213 /rpt_family="SVA"
repeat_region 42745, .43020 /rpt_family="Alu"
repeat_region complement(42763, .42966) /rpt_family="SVA"
repeat_region 43312, .43457 /rpt_family="Alu"
repeat_region complement(43963, .44217) /rpt_family="L1"
repeat_region complement(45296, .45533) /rpt_family="SVA"
repeat_region complement(45300, .45401) /rpt_family="L1"
repeat_region complement(45407, .45533) /rpt_family="L1"
repeat_region complement(45542, .45818) /rpt_family="Alu"
repeat_region 45598, .45770 /rpt_family="SVA"
repeat_region complement(45862, .46004) /rpt_family="L1"
repeat_region complement(45998, .46282) /rpt_family="Alu"
repeat_region 46022, .46260 /rpt_family="SVA"
repeat_region complement(46692, .47583) /rpt_family="Alu"
repeat_region 46885, .47215 /rpt_family="SVA"
repeat_region 47363, .47571 /rpt_family="SVA"
repeat_region complement(47665, .47773) /rpt_family="Alu"
repeat_region complement(47792, .48077) /rpt_family="Alu"
repeat_region complement(48110, .48388) /rpt_family="L1"
repeat_region complement(48827, .49102) /rpt_family="Alu"
repeat_region 48883, .49110 /rpt_family="SVA"
repeat_region complement(49395, .49626) /rpt_family="Alu"
repeat_region 50184, .50437 /rpt_family="Alu"
```

Query Match 80.7% Score 4034.2 DB 9: Length 142092;
Best Local Similarity 91.9% Pred. No. 0;
Matches 4486; Conservative 0; Mismatches 28; Indels 365; Gaps 7;
QY 122 GCAACCCACATGCAACGATACCTATGTAAACCACTGCACCATCATATACCAT 181
|||||

Db 38764 GCAAACCCACCATGGCAGACGTATACCTATGTAAACAACCTGCA - CATCATGTATACCTAT 38706

QY 182 GTAACCAACCTGGTATCATTTCTGCACACGTATCCACAGCATTTAGTGTAAAAAAGTG 241

Db 38705 GTTACAAACCT - GTACATTTCTGCACACGTATCCGAAAC - TTAAAGTAAAAAAGTG 38648

QY 242 GTGTGTAGAAAAATACCTGGCAATTCGACGATAGTTAAAGGTTAGTATTCAGAGGA 301

Db 38647 GTGTGTAGAAAAATACCTGGCAATTCGACGATAGTTAAAGGTTAGTATTCAGAGGA 301

QY 302 GAGGCTGACAGAAAGGAGATGAGATGGGTTAAGACAAAGTCATATTTAAAT 361

Db 38587 GAGGCTGACAGAAAGGAGATGAGATGGGTTAAGACAAAGTCATATTTAAAT 361

QY 362 CAGGCTTCTGGAAGTTAGTATTCGCAAAAACACACATCTCATATCCCTGACATGCTGA 421

Db 38527 CAGGCTTCTGGAAGTTAGTATTCGCAAAAACACACATCTCATATCCCTGACATGCTGA 421

QY 422 TTAAATGCAACATCCCTTAAGGCAACAGATCTGAGAAAGATCCGAAACATG 481

Db 38467 TTAAATGCAACATCCCTTAAGGCAACAGATCTGAGAAAGATCCGAAACATG 481

QY 482 CCGGTTTAAATTAATTTGGGCGATTTTACGCGCTCTTAAGAACCAAGAACATGCTGCG 38348

Db 38407 CCGGTTTAAATTAATTTGGGCGATTTTACGCGCTCTTAAGAACCAAGAACATGCTGCG 38348

QY 542 CTAGAGCTCTGCTCTCTTCATTTGAACATTAACAAGAGATGTGATAGTACACCCACCA 601

Db 38347 CTAGAGCTCTGCTCTCTTCATTTGAACATTAACAAGAGATGTGATAGTACACCCACCA 601

QY 602 CTTCACACAGCTTAGAGAGAGCCCTTGATAGATTTGATGATTAATTAATTTATTTGAATCA 38288

Db 38287 CTTCACACAGCTTAGAGAGAGCCCTTGATAGATTTGATGATTAATTAATTTATTTGAATCA 38288

QY 662 CATGCTAGATTTTACACAGAGTGGCGCTGGGAGCTGGGAGATTTATCCCATTTTGAATCA 38228

Db 38227 CATGCTAGATTTTACACAGAGTGGCGCTGGGAGATTTATCCCATTTTGAATCA 38228

QY 722 GGGTGGCTGGAACCCAGACGATTAATCTCAGAGGCTGTCTGCAACCCCAACAGACTG 781

Db 38167 GGGTGGCTGGAACCCAGACGATTAATCTCAGAGGCTGTCTGCAACCCCAACAGACTG 781

QY 782 ACCCCCTTCACACAGCCCGGAGAGCTTCTCCATCTCTGTAAGAACTTACTGAT 841

Db 38107 ACCCCCTTCACACAGCCCGGAGAGCTTCTCCATCTCTGTAAGAACTTACTGAT 841

QY 842 GGGCCCTGCGACGCAATCAGACGCGAATTAAGTATGATCACCAGCAGCAATCAGAG 901

Db 38047 GGGCCCTGCGACGCAATCAGACGCGAATTAAGTATGATCACCAGCAGCAATCAGAG 901

QY 902 CTCCGTCGACATATGAGAAATTCGTATTTACTAAGGTTGATGAAATTTCCCAAGCA 961

Db 37987 CTCCGTCGACATATGAGAAATTCGTATTTACTAAGGTTGATGAAATTTCCCAAGCA 961

QY 962 CACCATCTTTTTCAGATTAAGGCGACAGCTGAGAGAGAGTGAATACTACCCGGGT 1021

Db 37927 CACCATCTTTTTCAGATTAAGGCGACAGCTGAGAGAGAGTGAATACTACCCGGGT 1021

QY 1022 CACCAACACAGGTTGCAAGGCTGTCAAAAGCCAGAGGCTAGTCAAGGCTCTT 1141

Db 37807 TTTCAATCTTTTCATGAGCCACAGGCTGTCAAAAGCCAGAGGCTAGTCAAGGCTCTT 1141

QY 1082 TTTCAATCTTTTCATGAGCCACAGGCTGTCAAAAGCCAGAGGCTAGTCAAGGCTCTT 1141

Db 37807 TTTCAATCTTTTCATGAGCCACAGGCTGTCAAAAGCCAGAGGCTAGTCAAGGCTCTT 1141

QY 1142 CTTTCTGAGAGTCTGTGCGACAGAAATTTGAAGTTCAGACAGCCCTTAACCCCAAC 1201

Db 37747 CTTTCTGAGAGTCTGTGCGACAGAAATTTGAAGTTCAGACAGCCCTTAACCCCAAC 1201

QY 1202 TCTCTGTGCAAGGCTCAAGGCTCAGAACTGTTGAGAGATCTTTAGGCTCTGAT 1261

Db 37687 TCTCTGTGCAAGGCTCAAGGCTCAGAACTGTTGAGAGATCTTTAGGCTCTGAT 1261

QY 37687 TCTCTGTGCAAGGCTCAAGGCTCAGAACTGTTGAGAGATCTTTAGGCTCTGAT 37628

QY 1262 TTTAGGCCATGCTAGAGGGGGTGTGCCCTAAATTCCAGCCCTGTCTCAGCCCAACAC 1321
Db 37627 TTTAGGGCCATGCTAGAGGGGGTGTGCCCTAAATTCCAGCCCTGTCTCAGCCCAACAC 37566
QY 1322 CCTCCAGAAGAAATTAGAGGGGCCATGCGCAGGGCGTCTTAGCCGTTCTTCTGACAG 1381
Db 37567 CCTCCAGAAGAAATTAGAGGGGCCATGCGCAGGGCGTCTTAGCCGTTCTTCTGACAG 37508
QY 1382 ATTACAAAGAGGATTAAGACAAGACATCTTTGTGAGAGTCTGTGGCTTAGGAGATCAAG 1441
Db 37507 ATTACAAAGAGGATTAAGACAAGACATCTTTGTGAGAGTCTGTGGCTTAGGAGATCAAG 37448
QY 1442 TGAGGGGGGCTCAGCACTCAGCTGGGGAGTCCAGACCTCTAAGAGTGGGCAGGGGCACTG 1501
Db 37447 TGAGGGGGGCTCAGCACTCAGCTGGGGAGTCCAGACCTCTAAGAGTGGGCAGGGGCACTG 37388
QY 1502 GCCACAGATGCCAGGAGTCCCAACCAAGCCCTAGTCCGACAGCCTTCTGTGGGATCATGCG 1561
Db 37387 GCCACAGATGCCAGGAGTCCCAACCAAGCCCTAGTCCGACAGCCTTCTGTGGGATCATGCG 37328
QY 1562 ACCCACTGGAACCCCACTGTGATGATCAAGAGTGGCCAGTGGGATCTGGGCTTG 1621
Db 37327 ACCCACTGGAACCCCACTGTGATGATCAAGAGTGGCCAGTGGGATCTGGGCTTG 37268
QY 1622 GAGGCTTTCAGGGTTGGATGGCCATCTTCCGTAATTTGTGTGGGAATGACACACAGCA 1681
Db 37267 GAGGCTTTCAGGGTTGGATGGCCATCTTCCGTAATTTGTGTGGGAATGACACACAGCA 37208
QY 1682 GCACATGGGCGAGGTGTGGGCACTGTGTGTCTGTGCAAAATGCCCTAGGTGGGAATGA 1741
Db 37207 GCACATGGGCGAGGTGTGGGCACTGTGTGTCTGTGCAAAATGCCCTAGGTGGGAATGA 1742
QY 1742 GCTTGGTGTGCATCAGAGGAGCGACAGCCAGTGTGGCTGCAGCAAAACACACAGGAA 1801
Db 37147 GCTTGGTGTGCATCAGAGGAGCGACAGCCAGTGTGGCTGCAGCAAAACACACAGGAA 37091
QY 1802 AGATGGAGGGGCGCATTAATCTAGTACAAATTAATTAATGAGCTCCCCCTAAAAAAG 1861
Db 37090 AGATGGAGGGGCGCATTAATCTAGTACAAATTAATTAATGAGCTCCCCCTAAAAAAG 1862
QY 1862 AAGGTCTCTTTCGATGAGAAAGAGGAGAGGGGGTTTTCCTTAAATATAAGG 1921
Db 37067 AAGGTCTCTTTCGATGAGAAAGAGGAGAGGGGGTTTTCCTTAAATATAAGG 1922
QY 1922 AAGAGCCGCCCTCAAAAAAATAAGGAGGAGGAGACCAAGACCCCTGGGTGTGT 37068
Db 37067 AAGAGCCGCCCTCAAAAAAATAAGGAGGAGGAGACCAAGACCCCTGGGTGTGT 37067
QY 1982 TCCAGGGGAGCTCGAACCTTTAGAGGAGCGTGGGAACCGCTGTATTCAAGCCT 2041
Db 37067 TCCAGGGGAGCTCGAACCTTTAGAGGAGCGTGGGAACCGCTGTATTCAAGCCT 2042
QY 2042 CGAGAGAAAAGAGACGGCGCCCAAAAAAATATCCCTCCGGGGCATTAAGAAATGTGCC 2101
Db 37067 TCTCTCAAAAAGATGAAAGAGGACGGGATGTATGTATTATTTTAAAACTCAG 2161
QY 2162 GTAGNNNNNNNNNTCCTTCAGTAAATTTTATTAGAGCGCTCTCAAGAACACAAAG 2221
Db 37067 GTAGNNNNNNNNNTCCTTCAGTAAATTTTATTAGAGCGCTCTCAAGAACACAAAG 2222
QY 2222 GAGCTTCATTCGTGAGAGGAACAAGGCAAGCAACAGCAGATATCTGTATTAATTCAA 2281
Db 37021 GAGCTTCATTCGTGAGAGGAACAAGGCAAGCAACAGCAGATATCTGTATTAATTCAA 2282
QY 2282 GTAGGATTAATGCTCTTAAGAAATATCAAGATGAGTGAAGACACAGACACCGGTGG 2341
Db 36961 GTAGGATTAATGCTCTTAAGAAATATCAAGATGAGTGAAGACACAGACACCGGTGG 36902


```
repeat_region complement(52700, .52979)
/rpt_family="Alu"
repeat_region 53281, .53568
/rpt_family="Alu"
repeat_region 54071, .54204
/rpt_family="Alu"
repeat_region 54506, .54793
/rpt_family="Alu"
repeat_region 54807, .55061
/rpt_family="Alu"
repeat_region complement(64533, .64826)
/rpt_family="Alu"
repeat_region complement(65245, .65783)
/rpt_family="Alu"
repeat_region complement(68465, .68732)
/rpt_family="Alu"
repeat_region complement(70191, .70476)
/rpt_family="Alu"
repeat_region 70560, .70692
/rpt_family="Alu"
repeat_region 72294, .72470
/rpt_family="Alu"
repeat_region complement(72800, .73095)
/rpt_family="Alu"
repeat_region 73424, .73695
/rpt_family="Alu"
repeat_region 73784, .74071
/rpt_family="Alu"
repeat_region 74652, .74935
/rpt_family="Alu"
repeat_region complement(75265, .75514)
/rpt_family="Alu"
repeat_region complement(76086, .76353)
/rpt_family="Alu"
repeat_region complement(78038, .78312)
/rpt_family="Alu"
repeat_region 78513, .78734
/rpt_family="Alu"
repeat_region 78866, .79180
/rpt_family="Alu"
repeat_region complement(79384, .79673)
/rpt_family="Alu"
repeat_region complement(81398, .81690)
/rpt_family="Alu"
repeat_region 82013, .82288
/rpt_family="Alu"
repeat_region complement(83419, .83600)
/rpt_family="Alu"
repeat_region complement(87320, .87624)
/rpt_family="Alu"
repeat_region complement(87710, .87987)
/rpt_family="Alu"
repeat_region complement(88186, .88472)
/rpt_family="Alu"
repeat_region 88822, .89049
/rpt_family="Alu"
repeat_region 89224, .89308
/rpt_family="Alu"
repeat_region 89729, .90487
/rpt_family="Alu"
repeat_region 90504, .90618
/rpt_family="Alu"
repeat_region complement(93313, .93952)
/rpt_family="Alu"
repeat_region 96938, .97517
/rpt_family="Alu"
repeat_region complement(10419, .104514)
/rpt_family="Alu"
repeat_region complement(109738, .109979)
/rpt_family="Alu"
repeat_region 110022, .110141
/rpt_family="Alu"
repeat_region complement(111091, .112016)
```

```
Query Match 80.7%; Score 4034.2; DB 9; Length 196080;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 4486; Conservative 0; Mismatches 28; Indels 365; Gaps 7;

repeat_region /rpt_family="Alu"
113210, .113419
/rpt_family="Al"
repeat_region complement(114951, .115005)
/rpt_family="MIR"
repeat_region 115685, .115791
/rpt_family="MIR"
repeat_region 116870, .117005
/rpt_family="MIR"
repeat_region complement(117877, .118172)
/rpt_family="Alu"
repeat_region 118352, .118600
/rpt_family="Alu"
repeat_region 118679, .118813
/rpt_family="Alu"
repeat_region /rpt_family="Alu"
complement(121233, .121520)
repeat_region 122668, .123071

122 GCACCCACCATGACACAGCTATGTAACCACTGCACCATCATGATACAT 181
|||||
178042 GCACCCACCATGACACAGCTATGTAACCACTGCACCATCATGATACAT 178100
|||||
182 GTACCAACCTGCTGATCTGACACAGCTATGTAACCACTGCACCATCATGAT 241
|||||
178101 GTACCAACCTGCTGATCTGACACAGCTATGTAACCACTGCACCATCATGAT 178158
|||||
242 GTGTGTAGAAAATACCTGCAATCTGACATAGTTAAAGCTTATGATCTTACAGAGAGA 301
|||||
178159 GTGTGTAGAAAATACCTGCAATCTGACATAGTTAAAGCTTATGATCTTACAGAGAGA 178218
|||||
302 GAGGTGACAGAAAGGAGGATGAGAGCTGAGGTTTAAAGCAAGCTATTTAAAT 361
|||||
178219 GAGGTGACAGAAAGGAGGATGAGAGCTGAGGTTTAAAGCAAGCTATTTAAAT 178278
|||||
362 CAGGCTCTGTGAACTTATGATCCCAAAACACATCTCATATCCCTGCAGTCTTGA 421
|||||
178279 CAGGCTCTGTGAACTTATGATCCCAAAACACATCTCATATCCCTGCAGTCTTGA 178338
|||||
422 TTAATAATGCAATCCCTTAAGGCCACAGACTCAATCTGTGAGAAAGATCCAGAAAATG 481
|||||
178339 TTAATAATGCAATCCCTTAAGGCCACAGACTCAATCTGTGAGAAAGATCCAGAAAATG 178398
|||||
482 CCCGTTTAAATTAACATTTGGGCGATCTTACGGCTCTTAAAGACCAAGACCACTGCTGC 541
|||||
178399 CCCGTTTAAATTAACATTTGGGCGATCTTACGGCTCTTAAAGACCAAGACCACTGCTGC 178458
|||||
542 CTAGAGCTCTGCTCTCTTCAATTAACAAATACAGAGAGTGTAGTAGACACCCACCA 601
|||||
178459 CTAGAGCTCTGCTCTCTTCAATTAACAAATACAGAGAGTGTAGTAGACACCCACCA 178518
|||||
602 CTTCACACAGCTTGAAGAGCCCTGAGTATGATGATGATTAATTAATTTATGATCA 661
|||||
178519 CTTCACACAGCTTGAAGAGCCCTGAGTATGATGATGATTAATTAATTTATGATCA 178578
|||||
662 CATGCTGAGATTTTACACAGCTCCCGTGGGATCTGGGCAATTTATCCATTTGACAT 721
|||||
178579 CATGCTGAGATTTTACACAGCTCCCGTGGGATCTGGGCAATTTATCCATTTGACAT 178638
|||||
722 GCGTGTGAGAGCGACAGCAATTAATCTCAGGCGCTTCTGCAACCCCCACAGACTG 781
|||||
178639 GCGTGTGAGAGCGACAGCAATTAATCTCAGGCGCTTCTGCAACCCCCACAGACTG 178698
|||||
782 ACCCCCTCCACAGCCCGGAGGCTTCTCTTCATCTCTCTGAAGCAATTTACTGAT 841
|||||
178699 ACCCCCTCCACAGCCCGGAGGCTTCTCTTCATCTCTCTGAAGCAATTTACTGAT 178758
|||||
842 GGGCCGCGAGCATCAACGCAATTAATGATGATGATCAACGACGCAATTCAGAG 901
|||||
```

[illegible]

| | | | |
|-----------------|---|--|--------|
| Db | 181697 | GCTAAATTTTTTTTTTTTCTGATTATTTAGTAGACAGGGTTTCATCATGTGGCCAGG | 181756 |
| OY | 4202 | CNGGTGTTAACCCCAGCACTCAAGATGCACCACACTCGGCCCTCCCAAAGTCGGGA | 4261 |
| Db | 181757 | CTGGCTTTAACCCTCAGACCTCAAGATGCACCACACTCGGCCCTCCCAAAGTCGGGA | 181816 |
| OY | 4262 | TTACAGGTGTCAAGCCACATGCACAGCCCCACATGTTATTTTAAAAATATTTTAAA | 4321 |
| Db | 181817 | TTACAGGTGTCAAGCCACATGCACAGCCCCACATGTTATTTTAAAAATATTTTAAA | 181876 |
| OY | 4322 | TTAAAAATGTTTATCTTAAGGCCAGTTACAGTGATCTCGCGTCTGTAAATCCAGCACTTGGAG | 4381 |
| Db | 181877 | TTAAAAATGTTTATCTTAAGGCCAGTTACAGTGATCTCGCGTCTGTAAATCCAGCACTTGGAG | 181936 |
| OY | 4382 | GGGCCAAGGTGGGGGGATACATTGACCTGTGGAGTTTCAGGCTGGGCAACAATAGTAGACCC | 4441 |
| Db | 181937 | GGGCCAAGGTGGGGGGATACATTGACCTGTGGAGTTTCAGGCTGGGCAACAATAGTAGACCC | 181996 |
| OY | 4442 | CCGTCCTACCAAAAAATTTAAAAAATTTAGCTGGGAGTGTGGGCACTTTGCCCTGGTCCCA | 4501 |
| Db | 181997 | CCGTCCTACCAAAAAATTTAAAAAATTTAGCTGGGAGTGTGGGCACTTTGCCCTGGTCCCA | 182055 |
| OY | 4502 | GCTACTTGGGAAGCTGAGGTGTGGGGATGGCTTAACCTGTGAGGTGCAGAGCTGCAGTGA | 4561 |
| Db | 182057 | GCTACTTGGGAAGCTGAGGTGTGGGGATGGCTTAACCTGTGAGGTGCAGAGCTGCAGTGA | 182116 |
| OY | 4562 | GCTATGATCACACCACTGCACCTTCAGCTGAGTGACAGGCTATCTCAAAAGCAAAACAAA | 4621 |
| Db | 182117 | GCTATGATCACACCACTGCACCTTCAGCTGAGTGAGCAGGCTATCTCAAAAGCAAAACAAA | 182177 |
| OY | 4622 | TATGTGTTATCTTAAACGGTAAGGTATATATCACAGAATATATGATATAGCATTTTAAATTGAA | 4681 |
| Db | 182177 | TATGTGTTATCTTAAACGGTAAGGTATATATCACAGAATATATGATATAGCATTTTAAATTGAA | 182236 |
| OY | 4682 | AAAGCATTTATGATTTATCATGATGATTTGTAATAATATCAAAATACATGAATAATCTTGTGTTCTTA | 4741 |
| Db | 182237 | AAAGCATTTATGATTTATCATGATGATTTGTAATAATATCAAAATACATGAATAATCTTGTGTTCTTA | 182296 |
| OY | 4742 | ATTAATGCTTAGCAACAAGGCACATTTGGTTTAACTAGGGCACCAGAGTACTTTAAAAAAA | 4801 |
| Db | 182297 | ATTAATGCTTAGCAACAAGGCACATTTGGTTTAACTAGGGCACCAGAGTACTTTAAAAAAA | 182356 |
| OY | 4802 | GTTAGGGCCAGCCACAGGGGCTCACACCTGTATATCCACACACTTTGGGAGGCCAAGGCAG | 4861 |
| Db | 182357 | GTTAGGGCCAGCCACAGGGGCTCACACCTGTATATCCACACACTTTGGGAGGCCAAGGCAG | 182416 |
| OY | 4862 | GAGGATCACTTTAGCCCAAGGAGTTTAGGACCTGTGAGCAACAATAGGAGATCTGATCTTGTG | 4921 |
| Db | 182417 | GAGGATCACTTTAGCCCAAGGAGTTTAGGACCTGTGAGCAACAATAGGAGATCTGATCTTGTG | 182476 |
| OY | 4922 | CTCTATTAATAAATTTAAAAAATTTGGCTAGGCCCTTTGGCTTTAACCCGTAATCCACAGCT | 4981 |
| Db | 182477 | CTCTATTAATAAATTTAAAAAATTTGGCTAGGCCCTTTGGCTTTAACCCGTAATCCACAGCT | 182536 |
| OY | 4982 | TTGGGAGGCCGAGGCGGGT 5000 | |
| Db | 182537 | TTGGGAGGCCGAGGCGGGT 182555 | |
| RESULT 4 | | | |
| LOCUS | AC087451/c | | |
| DEFINITION | Homo sapiens chromosome 11 clone RP11-810P12 map 11, WORKING DRAFT SEQUENCE, 7 unordered pieces. | | |
| ACCESSION | AC087451 | | |
| VERSION | AC087451.2 GI:14595836 | | |
| KEYWORDS | HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP. | | |
| SOURCE | Homo sapiens. | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | |
| REFERENCE | 1 (bases 1 to 163915) | | |

QY 1521 TCCACAGCCTAGTGGCAGAGCCTGCTGGATCATGAGACCACCTGGAAACCACC 1580
| | | | |
Db 16315 TCCACAGCCTAGTGGCAGAGCCTGCTGGATCATGAGACCACCTGGAAACCACC 163256
| | | | |
QY 1581 TGTGAGTACAGAGGTGCCCCAGGTGAGTGGGCTTGGAGGCTTGAAGGTTGGAT 1640
| | | | |
Db 16325 TGTGAGTACAGAGGTGCCCCAGGTGAGTGGGCTTGGAGGCTTGAAGGTTGGAT 163196
| | | | |
QY 1641 GGGCATCTGCTATTTGTGTGGATATGACACACAGCAGACATGCGAGGTGTGT 1700
| | | | |
Db 163195 GGGCATCTGCTATTTGTGTGGATATGACACACAGCAGACATGCGAGGTGTGT 163136
| | | | |
QY 1701 GGCACCTGTGTCTGTGCAAAATGCTGAGGTGGAAATGAGTGGTGTGATGAGAG 1760
| | | | |
Db 163135 GGCACCTGTGTCTGTGCAAAATGCTGAGGTGGAAATGAGTGGTGTGATGAGAG 163078
| | | | |
QY 1761 CGACAGCAGCAGCTGTGGCTGACAGAAAACACAGAGGAAATGAGGGGATCAA 1820
| | | | |
Db 163077 -CACAGCAGCAGCTGTGGCTGACAGAAAACACAGAGGAAATGAGGGGATCAA 163019
| | | | |
QY 1821 TCACGTGACAAAATTTATATAGAGTCCCCCTAAAAAAGAGTCTCTCTTTGAT 1880
| | | | |
Db 163018 TCAC----- 163015
| | | | |
QY 1881 GAAGAAGGAGAGAGGGGCTTGTCTTAAATATAGAGAGAGCCGCCCTCAAAA 1940
| | | | |
Db 163014 ----- 163015
| | | | |
QY 1941 ATAAAGGAGAGAGAGCCCAAGACCCGTTGTGTGTTTCCAGGGGGAGCTGAAAC 2000
| | | | |
Db 163014 ----- 163015
| | | | |
QY 2001 CTTTAAAGGAGAGGCTGGAGAACCGCTGATTACAGCCTCTGAGAGAAAAGAGCGCC 2060
| | | | |
Db 163014 ----- 163015
| | | | |
QY 2061 GCCCAAAAATATCCCTCCCGGGCGATAGAAATGTGCTCTCAAAAAGATGAMGA 2120
| | | | |
Db 163014 ----- 163015
| | | | |
QY 2121 GGAAGCCGAGTGTATGTATGATATTTTAACTCCAGTAGNNNNNNNNNGCTT 2180
| | | | |
Db 163014 -----TGCCT 163010
| | | | |
QY 2181 CAGTAAATTTTATGAGCGCTTCTAGAGAACAAAGAGAGCTTCCATTCTAGAG 2240
| | | | |
Db 163009 CAGTAAATTTTATGAGCGCTTCTAGAGAACAAAGAGAGCTTCCATTCTAGAG 162950
| | | | |
QY 2241 GAAACAGCAGGAAACAGGAGATATCTGTATATTTCAAGTAGTATAGTCTCT 2300
| | | | |
Db 162949 GAAACAGCAGGAAACAGGAGATATCTGTATATTTCAAGTAGTATAGTCTCT 162890
| | | | |
QY 2301 AGAAATATCAAGCAAGGTGAGAGACACAGAGCAGCGGTGGAGTGGGCTCTATTCCA 2360
| | | | |
Db 162889 AGAAATATCAAGCAAGGTGAGAGACACAGAGCAGCGGTGGAGTGGGCTCTATTCCA 162830
| | | | |
QY 2361 GGTGGATGTTGGAGACATCTTCTTAAAGGAACTGAGTGGGAGGAAACATGAG 2420
| | | | |
Db 162829 GGTGGATGTTGGAGACATCTTCTTAAAGGAACTGAGTGGGAGGAAACATGAG 162770
| | | | |
QY 2421 GTATCTCAAGAGAGCTTCTTCCAGGAGAAAGATCAGCAGGTGGAAAGGCCCTGAGCC 2480
| | | | |
Db 162769 GTATCTCAAGAGAGCTTCTTCCAGGAGAAAGATCAGCAGGTGGAAAGGCCCTGAGCC 162710
| | | | |
QY 2481 ACCATTGAGTAAACATCTTGGAGATCTTACAGAGTATAGTTCATTATGAGAAATGGA 2540
| | | | |
Db 162709 ACCATTGAGTAAACATCTTGGAGATCTTACAGAGTATAGTTCATTATGAGAAATGGA 162650
| | | | |
QY 2541 ATATGAGTGTGAGAGAGGCTGCTGCTCTTCCATATCTTCTACACATAGGAGTGTGAG 2600
| | | | |
Db 162649 ATATGAGTGTGAGAGAGGCTGCTGCTCTTCCATATCTTCTACACATAGGAGTGTGAG 162590
| | | | |
QY 2601 AGAGCTTGGAGCTTACAGAACAAATGGGCTGAGAAACAGTCCCTAGCCAGAGAGACTGA 2660
| | | | |

Db 162589 AGAGCTTGGAGAGTAAAGAACAAAGATGGGCTGAGAAACATGCTGACCCAGAGACTGA 162530
| | | | |
QY 2661 GCTTATGTGTAGACATTCCTGCTGTACTGCTTGTGCTGTATATTTATTTA 2720
| | | | |
Db 162529 GCTTATGTGTAGACATTCCTGCTGTACTGCTTGTGCTGTATATTTATTTA 162470
| | | | |
QY 2721 TTTATGATCTTAAAGACAGATTTGCTTCTTAAAGGCTTGAAGGCTTGAAGGCTGA 2780
| | | | |
Db 162469 TTTATGATCTTAAAGACAGATTTGCTTCTTAAAGGCTTGAAGGCTTGAAGGCTGA 162410
| | | | |
QY 2781 TCTCAGCTCAGTCAACCTCAGCTCTGAGATCAAGGATTTCTGCTGAGCTGCT 2840
| | | | |
Db 162409 TCTCAGCTCAGTCAACCTCAGCTCTGAGATCAAGGATTTCTGCTGAGCTGCT 162350
| | | | |
QY 2841 GAGTACGTGGATTTACAGGACCCGACAGCCCTGAGATATTTTGTATTTTATGA 2900
| | | | |
Db 162349 GAGTACGTGGATTTACAGGACCCGACAGCCCTGAGATATTTTGTATTTTATGA 162290
| | | | |
QY 2901 GAGACAGGTTTACACATGTGGCCAGAGCTGTGCAACTCTGAGCTTAGTATGATCA 2960
| | | | |
Db 162289 GAGACAGGTTTACACATGTGGCCAGAGCTGTGCAACTCTGAGCTTAGTATGATCA 162230
| | | | |
QY 2961 CCTGCTGAGCTTCCAAAGTGGGATTTATAGGATGAGCAGCTGCGCCAGTATTA 3020
| | | | |
Db 162229 CCTGCTGAGCTTCCAAAGTGGGATTTATAGGATGAGCAGCTGCGCCAGTATTA 162170
| | | | |
QY 3021 TAGAAAGTTAAAGGACATGCGCAATGACAGCGCTATCTAGCTTCCCTCCAAAGCA 3080
| | | | |
Db 162169 TAGAAAGTTAAAGGACATGCGCAATGACAGCGCTATCTAGCTTCCCTCCAAAGCA 162110
| | | | |
QY 3081 AGGCGAGCTGCTGGGCTACATTTCTGCTTCTTACTTCCAAAAGCAGTACAGCTGCG 3140
| | | | |
Db 162109 AGGCGAGCTGCTGGGCTACATTTCTGCTTCTTACTTCCAAAAGCAGTACAGCTGCG 162050
| | | | |
QY 3141 AGGCGCTTGGAGACCTTCCACACCTCTAGGCTGCTTGGAGAGTGGAGTGGAGTCCAG 3200
| | | | |
Db 162049 AGGCGCTTGGAGACCTTCCACACCTCTAGGCTGCTTGGAGAGTGGAGTGGAGTCCAG 161990
| | | | |
QY 3201 AGCAGGGAAGGCTCTGACAGAGCTCTGACAGAGGCTCTGATCCCTACAAACCCCAATC 3260
| | | | |
Db 161989 AGCAGGGAAGGCTCTGACAGAGCTCTGACAGAGGCTCTGATCCCTACAAACCCCAATC 161930
| | | | |
QY 3261 GGTGCTCCTCTCTACAGAGCAACAGCCAGCTCTGAGAGCCAGCTGCTGAGGCTATGAC 3320
| | | | |
Db 161929 GGTGCTCCTCTCTACAGAGCAACAGCCAGCTCTGAGAGCCAGCTGCTGAGGCTATGAC 161870
| | | | |
QY 3321 ATCACTTACACAGCAAGTGGCTAATGCCGCTTAGGCTCTTCTCCGCTGCTGCTG 3380
| | | | |
Db 161869 ATCACTTACACAGCAAGTGGCTAATGCCGCTTAGGCTCTTCTCCGCTGCTGCTG 161810
| | | | |
QY 3381 TGTGGGGGGGACATCTACAAAGCTCTTATAGGAGAGTTCCTAATCTTCTGCTGCTG 3440
| | | | |
Db 161809 TGTGGGGGGGACATCTACAAAGCTCTTATAGGAGAGTTCCTAATCTTCTGCTGCTG 161750
| | | | |
QY 3441 TACTACATCATCGCTTATTTATAGTAAAGCTGAGAGGCTGGGCGGGGGCTGGG 3500
| | | | |
Db 161749 TACTACATCATCGCTTATTTATAGTAAAGCTGAGAGGCTGGGCGGGGGCTGGG 161690
| | | | |
QY 3501 AAGGATGTGGCTGGGCTGGAGCTGGGAGCTCTGGGGCTTCCAGGCAAGCTAGGGC 3560
| | | | |
Db 161689 AAGGATGTGGCTGGG-----GCTGGAGCTCTTGGGGCTTCCAGGCTAGGGG 161637
| | | | |
QY 3561 CCAAGTACAGCTTCACTACAAACATTAAGCTGGGCTCTTCCAGGCTCTGGGCACTGGA 3620
| | | | |
Db 161636 CCAAGTACAGCTTCACTACAAACATTAAGCTGGGCTCTTCCAGGCTCTGGGCACTGGA 161577
| | | | |
QY 3621 GCTGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGAA 3680
| | | | |
Db 161576 GCTGAGGCTG-GGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGAA 161518
| | | | |
QY 3681 GCCAGCCAGCAGAGGTTTATAGCAGCTTCTCTCAACCCAGAGAGAGCCCTGAGAGCCGA 3740
| | | | |

```

Db 161517 GCCAGGCGCAGAGGGTTTATGACCACTTCTCCAAACCCAGAGAGGCCCTGAGGCCA 161458
OY 3741 GGGTTGTGCTGGCCCACTCTACTGCGCTGTTTCTGTAATCCACAGACTCTATAGCC 3800
Db 161457 GGGTTGTGCTGGCCCACTCTACTGCGCTGTTTCTGTAATCCACAGACTCTATAGCC 3800
OY 3801 CCACATAGTACATTTAAAAA--AGAGAGAGAGAGAGAGAGAGAGAGAGATGAGATGCA 161398
Db 161397 CCACATAGTACATTTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGATGAGATGCA 3858
OY 3859 CTGTTGTGCTGGCCCACTCTACTGCGCTGTTTCTGTAATCCACAGACTCTATAGCC 3918
Db 161337 CTGTTGTGCTGGCCCACTCTACTGCGCTGTTTCTGTAATCCACAGACTCTATAGCC 3918
OY 3919 CCAAGGGGCTGGAGTTTACAGGTGAGCTGACCTGACCTGCAATCCCTGCTTATAGCCCTC 161278
Db 161277 CCAAGGGGCTGGAGTTTACAGGTGAGCTGACCTGACCTGCAATCCCTGCTTATAGCCCTC 161278
OY 3978 TTTTGTGTTTGTGTTTGTGAGACAGAGGTTTCACTCCATCCAGCCAGCTGGAGTGCAGTG 161218
Db 161217 TTTTGTGTTTGTGTTTGTGAGACAGAGGTTTCACTCCATCCAGCCAGCTGGAGTGCAGTG 161158
OY 4038 GGGGCAATCTTGGCTACAGTACCTCTGCTCCAGGTGCAAGCATCTCTGCTTATAGCCCT 161157
Db 161157 GGGGCAATCTTGGCTACAGTACCTCTGCTCCAGGTGCAAGCATCTCTGCTTATAGCCCT 161098
OY 4098 GCTCTGAGTACGTGAGTATTAAGCAGACACACACACACCTGCTGCTTATAGCCCTTAT 161097
Db 161097 GCTCTGAGTACGTGAGTATTAAGCAGACACACACACACCTGCTGCTTATAGCCCTTAT 161038
OY 4158 TTTCTATTTTATTTATAGAGACAGAGGTTTCTATCATGTTGGCGAGGCTTGAACCCCT 4217
Db 161037 TTTCTATTTTATTTATAGAGACAGAGGTTTCTATCATGTTGGCGAGGCTTGAACCCCT 160978
OY 4218 GACCTCAAGTATCCACACACCTGCTCCCAAGTCTGGAGTATACAGGTGTCAGCA 4277
Db 160977 GACCTCAAGTATCCACACACCTGCTCCCAAGTCTGGAGTATACAGGTGTCAGCA 160918
OY 4278 CCATGACACAGCCCACTGATGATTTTAAATTTTATTTTATTTAAATGTTTATCTA 4337
Db 160917 CCATGACACAGCCCACTGATGATTTTAAATTTTATTTTATTTAAATGTTTATCTA 160858
OY 4338 AGGCGATAGCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4397
Db 160857 AGGCGATAGCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4398
OY 4398 ATCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 160798
Db 160797 ATCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4457
OY 4458 TTTTAAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 160738
Db 160737 TTTTAAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4517
OY 4518 AGCTGTGGGATGCTGGAAGCTGTGAGGTGAGCTGAGTGTGAGCTATGATCAGACAC 4577
Db 160677 AGCTGTGGGATGCTGGAAGCTGTGAGGTGAGCTGAGTGTGAGCTATGATCAGACAC 4577
OY 4578 TGCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 160618
Db 160617 TGCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4637
OY 4638 GGTAAAGTATATACAGATATATATATATATATATATATATATATATATATATATATAT 160558
Db 160557 AATTAATATATATATATATATATATATATATATATATATATATATATATATATAT 4697
OY 4698 CATGATTTGTTTATATATATATATATATATATATATATATATATATATATATATATAT 160498
Db 160497 CATGATTTGTTTATATATATATATATATATATATATATATATATATATATATATAT 4757
OY 4758 GGCACATTTGTTTATATATATATATATATATATATATATATATATATATATATATAT 160438
Db 160437 GGCACATTTGTTTATATATATATATATATATATATATATATATATATATATATAT 160378

```

```

OY 4818 GGGGCTACACCTGTAATCCAGCAGCTTTGGAGGCCCAAGCAGAGATGATCTTGAACC 4877
Db 160377 GGGGCTACACCTGTAATCCAGCAGCTTTGGAGGCCCAAGCAGAGATGATCTTGAACC 160318
OY 4878 CAGGAGTTTGGAGCCTAGACACATATAGGAGATCTTGTCTCTATATATATATATAT 4937
Db 160317 CAGGAGTTTGGAGCCTAGACACATATAGGAGATCTTGTCTCTATATATATATATAT 160238
OY 4938 AAAATTTGGTACGCTTTTGTCTTACACCCGTAATCCAGCAGCTTTGGAGGCCGAG 4997
Db 160257 AAAATTTGGTACGCTTTTGTCTTACACCCGTAATCCAGCAGCTTTGGAGGCCG 160198
OY 4998 GGT 5000
Db 160197 GGT 160195

RESULT 5
AP003733 16867 bp DNA linear PRI 08-DEC-2001
LOCUS AP003733
DEFINITION Homo sapiens genomic DNA, chromosome 11q, clone: RP11-810P12, complete sequence.
ACCESSION AP003733
VERSION AP003733.4 GI:17426128
KEYWORDS HTG.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Matanabe, H. and Sakaki, Y. Published only in Database (2001) 2 (bases 1 to 166867)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Matanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (11-JUN-2001) Masahiro Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suenho-chou, Tsukuba, Ibaraki, 305-3858, Japan (E-mail: hattori@gs.c.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT On Dec 7, 2001 this sequence version replaced g1:17026124.
FEATURES
source
1. 166867
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-810P12"
BASE COUNT 44705 a 38465 c 39268 g 44429 t
ORIGIN
Query Match 65.18; Score 3254.6; DB 9; Length 166867;
Best Local Similarity 90.58; Pred. No. 0;
Matches 3694; Conservative 0; Mismatches 24; Indels 365; Gaps 6;
OY 921 GAATTCGTGCTATTTTACTAGGATGATGAATTTCCCAAGCAACACATCTTTTCAGATAA 980
Db 1 GAATTCGTGCTATTTTACTAGGATGATGAATTTCCCAAGCAACACATCTTTTCAGATAA 980
OY 981 GGGCAGTGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1040
Db 61 GGGCAGTGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
OY 1041 GGGTGGAGCAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1100
Db 121 GGGTGGAGCAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
OY 1101 AAGAGGCTGTCAAAAGCCAGGCTGTGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1160

```

| | | | | | | | |
|---|---|--|------|---|-----------|---------------------------------|-----|
| D | b | | 181 | ACAGGGCTGTCAAAAGACC | CCCGGCTTA | GTCAGAGGCTCTCCTTCTTGAGAGTTCTCTG | 240 |
| O | y | | 1161 | GCACAGAAGTTGAAGTCTACAGACAGGCCCTTAACCCCACAATCTCTGCAAGGCTTAG | 1220 | | |
| D | b | | 241 | GCACAGAAGTTGAAGTCTACAGACAGGCCCTTAACCCCACAATCTCTGCAAGGCTTAG | 300 | | |
| O | y | | 1221 | GGGTCAAGAACACTGGTGTGAGNCAAGATCTTTAGCCTGTGAAATTTAAGGGCATGGTAGAG | 1280 | | |
| D | b | | 301 | GGGTCAAGAACACTGGTGTGAGNCAAGATCTTTAGCCTGTGAAATTTAAGGGCATGGTAGAG | 360 | | |
| O | y | | 1281 | GGGTGTGGCCCTAAATTCCAGGCCCTGGTCTCAGGCCCAACCCCTCCAGAAGAAATTAAGA | 1340 | | |
| D | b | | 361 | GGGTGTGGCCCTAAATTCCAGGCCCTGGTCTCAGGCCCAACCCCTCCAGAAGAAATTAAGA | 420 | | |
| O | y | | 1341 | GGGGCCATATGCCACAGCTGTGCTAAGCCGTCTTCTGAGCAGATTACAAGAAAGGACTAAG | 1400 | | |
| D | b | | 421 | GGGGCCATATGCCACAGCTGTGCTAAGCCGTCTTCTGAGCAGATTACAAGAAAGGACTAAG | 480 | | |
| O | y | | 1401 | ACAAGAGACTCTTTTGGAGAGTCTGGGCTTAAGGAGTGAAGTGAAGGGGGGTCAAGACTC | 1460 | | |
| D | b | | 481 | ACAAGAGACTCTTTTGGAGAGTCTGGGCTTAAGGAGTGAAGTGAAGGGGGGTCAAGACTC | 540 | | |
| O | y | | 1461 | ACGTGGGACAGTGTCCAGCCTCTTAAGAGTGGGACAGGGGCACTGGCCAAGAGTCCCCAGGAG | 1520 | | |
| D | b | | 541 | ACGTGGGACAGTGTCCAGCCTCTTAAGAGTGGGACAGGGGCACTGGCCAAGAGTCCCCAGGAG | 600 | | |
| O | y | | 1521 | TCCCAACAGCCTAAGTGGGACAGCCTTCTGTGGAGTACATCGAACCCACTCTGGAAACCCACC | 1580 | | |
| D | b | | 601 | TCCCAACAGCCTAAGTGGGACAGCCTTCTGTGGAGTACATCGAACCCACTCTGGAAACCCACC | 660 | | |
| O | y | | 1581 | TGTGAGTACAAAGTGCCCCAGGTGACHTGGGCTTGAGGCTTTCAGAGCTTTCAGGTTGGAT | 1640 | | |
| D | b | | 661 | TGTGAGTACAAAGTGCCCCAGGTGACHTGGGCTTGAGGCTTTCAGAGCTTTCAGGTTGGAT | 720 | | |
| O | y | | 1641 | GGCCATCTTGGGTAATTTGTGGGATATGCACACACAGGCGACATGCGCAGGTGTGTG | 1700 | | |
| D | b | | 721 | GGCCATCTTGGGTAATTTGTGGGATATGCACACACAGGCGACATGCGCAGGTGTGTG | 780 | | |
| O | y | | 1701 | GGCACCTTGTGTCTGTGCAAAATCCCTGAGAGTGGGATGAGCTTGTGTGCATCAAGAG | 1760 | | |
| D | b | | 781 | GGCACCTTGTGTCTGTGCAAAATCCCTGAGAGTGGGATGAGCTTGTGTGCATCAAGAG | 840 | | |
| O | y | | 1761 | CGACAGCCAGGACAGTGGGCTGCGACGAACAAACACACAGGAAAGATGAGAGGGGATCAA | 1820 | | |
| D | b | | 839 | -CACAGCCAGGACAGTGGGCTGCGACGAACAAACACAGGAAAGATGAGAGGGGATCAA | 897 | | |
| O | y | | 1821 | TCACGTACAAAATTAATTTATAGAGCTCCCTTAAGAAAAAGAGTCTCTTCTTGATA | 1880 | | |
| D | b | | 898 | TCAC----- | 901 | | |
| O | y | | 1881 | GAAAGAGGGAAGAGAGGGGTTGTCTTAATTAATATAGGAGAGAGCCGCCCTCAAAA | 1940 | | |
| D | b | | 902 | ----- | 901 | | |
| O | y | | 1941 | ATAAGGAGGAGAGACCCCAAGACCCTGGGTGTGTGTTTTCCAGGGGAGTCTGAACC | 2000 | | |
| D | b | | 902 | ----- | 901 | | |
| O | y | | 2001 | CTTTAGAGGAGCGTGGGAGAACCGCTGTATTCAAGCCCTCTGAGAGAAAAAGAGCGGCC | 2060 | | |
| D | b | | 902 | ----- | 901 | | |
| O | y | | 2061 | GCCCAAAAATATCCCTCCCGGGGAGATAAATAATGGTGGGCTCTCAAAAAAGATGAGA | 2120 | | |
| D | b | | 902 | ----- | 901 | | |
| O | y | | 2121 | GGAAGCGGAGTGTATGTGTGATATTTTAAACCTCAGTAGNNNNNNNNNTGCTT | 2180 | | |
| D | b | | 902 | -----TGCTT | 906 | | |
| O | y | | 2181 | CAGTAAATTTTATATGAGCGCCTTCTACAGAAACACAGAGAGCTTCCATTTCTGAGAG | 2240 | | |
| D | b | | 907 | CAGTAAATTTTATATGAGCGCCTTCTACAGAAACACAGAGAGCTTCCATTTCTGAGAG | 966 | | |

| | | | |
|----|------|---|------|
| QY | 2241 | GAAGACGGCAGGAAAGGACAGTATCCGTCTAATATTCGAATAGTACTGATAGTGCCTCT | 2300 |
| Db | 967 | GAAGACGGCAGGAAAGGACAGTATCCGTCTAATATTCGAATAGTACTGATAGTGCCTCT | 1028 |
| QY | 2301 | AGAAATATACAGCMAAGGTAGAGAGACACAGACCGGTGGGTGAGTGGGACTCTATTTCGA | 2360 |
| Db | 1027 | AGAAATATACAGCMAAGGTAGAGAGACACAGACCGGTGGGTGAGTGGGAAACCATGTGAG | 1146 |
| QY | 2361 | GGATGGATGGTGGGAAACATCCTTTCTAAAGGAAACCTGAGATGGGAAAGCAACATGAC | 2420 |
| Db | 1087 | GGTGGATGGTGGGAAACATCCTTTCTAAAGGAAACCTGAGATGGGAAAGCAACATGAC | 1146 |
| QY | 2421 | GTATCTCAGGAAGAGCTTCTCCACGAGCAGGAAGATCAGCAGGTGGAAAGGCCCTGAGACC | 2480 |
| Db | 1147 | GTATCTCAGGAAGAGCTTCTCCAGGACAGGAAGATCAGCAGGTGGAAAGGCCCTGAGACC | 1206 |
| QY | 2481 | ACCATTCAGTAAACATCATTTAGAGATCCTTCACAGCTAGAGTTCATATATGGGAATGGGA | 2540 |
| Db | 1207 | ACCATTCAGTAAACATCATTTAGAGATCCTTCACAGCTAGAGTTCATATATGGGAATGGGA | 1266 |
| QY | 2541 | ATATGGTGGTGAACAGGCGCTGCTGTCCTTCATACCTCTCAGACATAGGGTGGTTGAG | 2600 |
| Db | 1267 | ATATGGTGGTGAACAGGCGCTGCTGTCCTTCATACCTCTCAGACATAGGGTGGTTGAG | 1326 |
| QY | 2601 | AGAGCTTGGGACTCTACGAAACAAGATGGGCTGAGAAACATCCGTAACCCCAAGAACCTGA | 2660 |
| Db | 1327 | AGAGCTTGGGACTCTACGAAACAAGATGGGCTGAGAAACATCCGTAACCCCAAGAACCTGA | 1386 |
| QY | 2661 | GCTTATGTGTAGACATTCCTGCTGCTTACTGCTTGTCTGTTATATTAATTAATTAATTA | 2720 |
| Db | 1387 | GCTTATGTGTAGACATTCCTGCTGCTTACTGCTTGTCTGTTATATTAATTAATTAATTA | 1446 |
| QY | 2721 | TTTATTTGATCTTAACACAGATTTTGGCTTTACCCAGGCTTGAGGCAATGGCGCTGA | 2780 |
| Db | 1447 | TTTATTTGATCTTAACACAGATTTTGGCTTTACCCAGGCTTGAGGCAATGGCGCTGA | 1506 |
| QY | 2781 | TCCTCAGCTACCTGCAACCTCCACTCTGGGATCAAGCATTCCTGCTCAGCCTCTCT | 2840 |
| Db | 1507 | TCCTCAGCTACCTGCAACCTCCACTCTGGGATCAAGCATTCCTGCTCAGCCTCTCT | 1566 |
| QY | 2841 | GAGTAGCTGGGATTTACAGGACACCGGACACAGCCTGGATTAATTTTGTATTTTATTA | 2900 |
| Db | 1567 | GAGTAGCTGGGATTTACAGGACACCGGACACAGCCTGGATTAATTTTGTATTTTATTA | 1526 |
| QY | 2901 | GAGAAAGGCTTACACATGTTGGCCAGGCGTGGTCTCGAATCTGACCTTAAGGTATCCA | 2960 |
| Db | 1627 | GAGAAAGGCTTACACATGTTGGCCAGGCGTGGTCTCGAATCTGACCTTAAGGTATCCA | 1686 |
| QY | 2961 | CCTGCGCTGACATCCCAAAAGTGTGGGATTTAAGCATGAGCCACTCGGCCAGATGTTA | 3020 |
| Db | 1687 | CCTGCGCTGACATCCCAAAAGTGTGGGATTTAAGCATGAGCCACTCGGCCAGATGTTA | 1746 |
| QY | 3021 | TAGAAAGTTAAAGGCACATGGCAATGCAACAGCCTATCTACGCTTCCCTGCCAAAGCAA | 3080 |
| Db | 1747 | TAGAAAGTTAAAGGCACATGGCAATGCAACAGCCTATCTACGCTTCCCTGCCAAAGCAA | 1806 |
| QY | 3081 | AGGGCAGCCTCTGGGGCTACATTTCTCGATTTCTACTTCCAAAAGGACAGTCAACAATGGC | 3140 |
| Db | 1807 | AGGGCAGCCTCTGGGGCTACATTTCTCGATTTCTACTTCCAAAAGGACAGTCAACAATGGC | 1866 |
| QY | 3141 | AGGGCCTTGGAGACCACTTCAATCCACTCTTAGGGTCCCTATGGGAGAGTTGAGGTCCAG | 3200 |
| Db | 1867 | AGGGCCTTGGAGACCACTTCAATCCACTCTTAGGGTCCCTATGGGAGAGTTGAGGTCCAG | 1926 |
| QY | 3201 | AGCAGGGAAGGAGTCTCTACAGAGGCTCTACACAGGCGCTGATCCGTAACAACCCGCAATC | 3260 |
| Db | 1927 | AGCAGGGAAGGAGTCTCTACAGAGGCTCTACACAGGCGCTCTGATCCGTAACAACCCGCAATC | 1986 |
| QY | 3261 | GGTGTCCCTCTTACAGAGACCAAGCCACGCTGTGACAGCCCACTGCTGGCCATGACC | 3320 |
| Db | 1987 | GGTGTCCCTCTTACAGAGACCAAGCCACGCTGTGACAGCCCACTGCTGGCCATGACC | 2046 |

3321 ATCACTTACACAGCAAGTGGCTAATGCCGCTTAGGCTCTCTCCGCTCTGCTG 3380
DB ATCACTTACACAGCAAGTGGCTAATGCCGCTTAGGCTCTCTCCGCTCTGCTG 2106
3381 TGGCGGGGGGAGCAATCTACAGCTCTATATGAGGAGTTCCTAATCTCTGCTGCTG 3440
DB TGGCGGGGGGAGCAATCTACAGCTCTATATGAGGAGTTCCTAATCTCTGCTGCTG 2166
3441 TACTACATCATCCGCTTATTTATAGTAAAGCTGCGAGGCTGGGCGGGGGGCTGG 3500
DB TACTACATCATCCGCTTATTTATAGTAAAGCTGCGAGGCTGGGCGGGGGGCTGG 2226
3501 AAGGATGTGGCTGGGCTGGAGCTGGAGCTCTGCGGGGCTCCAGCCAGCTCAAGGC 3560
DB AAGGATGTGGCTGGGCTGGAGCTGGAGCTCTGCGGGGCTCCAGCCAGCTCAAGGC 2279
3561 CCAGTCCACCACTCCCTCAACACTAAGCTGGGCTCTGAGCAGCTCTGCGGCACTG 3620
DB CCAGTCCACCACTCCCTCAACACTAAGCTGGGCTCTGAGCAGCTCTGCGGCACTG 2339
3621 GCTGAGGCTGCGGCTGGGCTGGAGAGTAAAGAGTCACTGAGAGGCTGCTCA 3680
DB GCTGAGGCTGCGGCTGGGCTGGAGAGTAAAGAGTCACTGAGAGGCTGCTCA 2398
3681 GCGAGGCGCAGAGGCTTTAGCCACCTCTCCCAACCCAGAGGAGCCCTGGAGCCCA 3740
DB GCGAGGCGCAGAGGCTTTAGCCACCTCTCCCAACCCAGAGGAGCCCTGGAGCCCA 2458
3741 GCGTGTGTGCTGGCCCTCTACAGCTGCTGTTTACTGAAATCCACAGAGCTATAGGC 3800
DB GCGTGTGTGCTGGCCCTCTACAGCTGCTGTTTACTGAAATCCACAGAGCTATAGGC 2518
3801 CCACATGTATATTAAAAA - AGAGAGAGAGAGAGAGAGAGAGAGAGAGTCA 3858
DB CCACATGTATATTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGTCA 2578
3859 CTGTGTGTGCTGGGCTGCTCCAGCTCTAGGCTCAAGCAATCCCTGCTGAGCTG 3918
DB CTGTGTGTGCTGGGCTGCTCCAGCTCTAGGCTCAAGCAATCCCTGCTGAGCTG 2638
3919 CCAGAGGCGCTGGATTTACAGAGTGTGAGTCACTGACCTGACCAACATGATG 3977
DB CCAGAGGCGCTGGATTTACAGAGTGTGAGTCACTGACCTGACCAACATGATG 2698
3978 TTTTGT 4037
DB TTTTGT 2698
2699 TTTTGT 2758
4038 GGGCAATCTTGGCTCACATGTAACCTCTGCTCCAGGCTGCAAGGATCTCTGCTG 4097
DB GGGCAATCTTGGCTCACATGTAACCTCTGCTCCAGGCTGCAAGGATCTCTGCTG 2818
4098 GCGTCTGTGAGTGTGGAATTTAGGACACACACACACACACACACACACACAC 4157
DB GCGTCTGTGAGTGTGGAATTTAGGACACACACACACACACACACACACACAC 2819
2819 GCGTCTGTGAGTGTGGAATTTAGGACACACACACACACACACACACACACAC 4157
DB GCGTCTGTGAGTGTGGAATTTAGGACACACACACACACACACACACACACAC 2878
4158 TTTGTATTTTGTATAGAGAGAGGCTTCAATGTTGGCCAGGCTGCTTGAACCT 4217
DB TTTGTATTTTGTATAGAGAGAGGCTTCAATGTTGGCCAGGCTGCTTGAACCT 2879
2879 TTTGTATTTTGTATAGAGAGAGGCTTCAATGTTGGCCAGGCTGCTTGAACCT 2938
4218 GACCTCAAGTGTAC 4277
DB GACCTCAAGTGTAC 2939
2939 GACCTCAAGTGTAC 4278
4278 CCATGACACACACACATGTATATTTTAAATTTTAAATTTTAAATTTTAAATTT 4337
DB CCATGACACACACACATGTATATTTTAAATTTTAAATTTTAAATTTTAAATTT 3058
4338 AGGCAAGTGTACAGTGTGCTGTATATCCAGACACTTTGAGGGGCAAGTGGGG 4397
DB AGGCAAGTGTACAGTGTGCTGTATATCCAGACACTTTGAGGGGCAAGTGGGG 3059
3059 AGGCAAGTGTACAGTGTGCTGTATATCCAGACACTTTGAGGGGCAAGTGGGG 3118
4398 ATCACTTGTAGGCTGGAGTGTACAGTGTGGGCAATATGTAGAGCCCGCTCTAC 4457

DB 3119 ATCACTTGTAGGCTGGAGTGTACAGTGTGGGCAATATGTAGAGCCCGCTCTAC 3178
QY 4458 TTTAAAAAATTTACTGGGAGTGTGTCATTTGCTGTGTGCCAGCTACTTGGAGCTG 4517
DB 3179 TTTAAAAAATTTACTGGGAGTGTGTCATTTGCTGTGTGCCAGCTACTTGGAGCTG 3238
QY 4518 AGGTGTGGGAGTGTGAGGCTGTGAGCTGTCAGAGCTGCAAGTGTATGATCAGACAC 4577
DB 3239 AGGTGTGGGAGTGTGAGGCTGTGAGCTGTCAGAGCTGCAAGTGTATGATCAGACAC 3298
QY 4578 TGCATCTTACGCTGAGTGTACAGCTATCTCAAAAGCAAAATTAATTTATCTAAC 4637
DB 3299 TGCATCTTACGCTGAGTGTACAGCTATCTCAAAAGCAAAATTAATTTATCTAAC 3358
QY 4638 GGTAAAGTATATATACAGAAATATATGATATGATTTTAAATTTGAAAAAGCATTAATG 4697
DB 3359 AATTAATATATATATACAGAAATATATGATATGATTTTAAATTTGAAAAAGCATTAATG 3418
QY 4698 CATGATTTGTAAATATATATATATATGATTTTGTGTGTCTTATATGCTAGCAAA 4757
DB 3419 CATGATTTGTAAATATATATATATATGATTTTGTGTGTCTTATATGCTAGCAAA 3478
QY 4758 GGCACATTTGTTTGTATCTAGAGGCAAGGCTACTTTAAAAAAGTTAGGCGCAGCA 4817
DB 3479 GGCACATTTGTTTGTATCTAGAGGCAAGGCTACTTTAAAAAAGTTAGGCGCAGCA 3538
QY 4818 GGGGCTCACACCTTATATCCAGACCTTTGGAGGCGCAGGAGCATCACTTGAGCC 4877
DB 3539 GGGGCTCACACCTTATATCCAGACCTTTGGAGGCGCAGGAGCATCACTTGAGCC 3598
QY 4878 CAGAGTGTAGAGCTGTAGGCAATAGGAGATCTGTATCTATATATATATTA 4937
DB 3599 CAGAGTGTAGAGCTGTAGGCAATAGGAGATCTGTATCTATATATATATTA 3658
QY 4938 AAAATGGCTAGAGCCCTTGGCTTACACCCGTTATCCAGACCTTTGGAGGCGCAGG 4997
DB 3659 AAAATGGCTAGAGCCCTTGGCTTACACCCGTTATCCAGACCTTTGGAGGCGCAGG 3718
QY 4998 GGT 5000
DB 3719 GGT 3721

RESULT 6
AC084857
LOCUS
DEFINITION Homo sapiens chromosome 11, clone CTD-3231N5 map 11, WORKING DRAFT
AC084857
ACCESSION
VERSION AC084857.1 GI:11276215
KEYWORDS HTG, HTGS, PHASE1, HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 133683)
TITLE Homo sapiens chromosome 11, clone CTD-3231N5
JOURNAL
REFERENCE
AUTHORS 2 (bases 1 to 133683)
Anderson, S., Barna, N., Bustin, C., Lander, E., Abraham, H., Allen, N.,
Bouhassira, B., Brown, A., Burt, G., Campobasso, A., Castle, A.,
Chapel, Y., Colangelo, M., Collins, S., Collinmore, A., Cooke, P.,
Delellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
Fitzhugh, W., Gage, D., Galagan, J., Gardina, S., Ginde, S., Goyette, M.,
Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., LaRocque, M.,
Lamzates, R., Landers, T., Lehotzky, J., Levine, R., Liu, C., Liu, G.,
Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,
McPherson, R., Meldrum, J., Meneus, L., Mlivo, T., Mlivo, Y.,
Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,

TITLE
JOURNAL
COMMENT

O'Donnell, P., O'Neill, D., Oliveira, T.M., Oliver, J., Peterson, K.,
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Reback, M., Riley, R.,
Rogov, P., Rothman, D., Roy, A., Santos, R., Schuer, S., Severy, P.,
Sougen, C., Spencer, B., Stange, Thomas, N., Stojanovic, N.,
Strauss, N., Sudriaman, A., Talamas, J., Testaye, S., Theodore, J.,
Tirrell, A., Travers, M., Trigg, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wymann, D., Ye, W.J., Young, G., Zainoun, J.,
Zimmer, A. and Zody, M.

Submitted (22-NOV-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu

Project Information
Center project name: 3231_N5
Center clone name: 3231_N5

Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 126761 bases at least Q40
Consensus quality: 129952 bases at least Q30
Consensus quality: 131351 bases at least Q20
Insert size: 128000; agarose-IP
Insert size: 132483; sum-of-ctrls
Quality coverage: 6.1 in Q20 bases; sum-of-ctrls
Quality coverage: 5.9 in Q20 bases; sum-of-ctrls

* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1459: contig of 1459 bp in length
* 1460 1559: gap of 100 bp
* 1560 2520: contig of 961 bp in length
* 2521 2620: gap of 100 bp
* 2621 6334: contig of 3714 bp in length
* 6335 6434: gap of 100 bp
* 6435 11913: contig of 5479 bp in length
* 11914 12013: gap of 100 bp
* 12014 20756: contig of 8743 bp in length
* 20757 20856: gap of 100 bp
* 20857 29735: contig of 8879 bp in length
* 29736 29835: gap of 100 bp
* 29836 38645: contig of 8810 bp in length
* 38646 38745: gap of 100 bp
* 38746 52193: contig of 13448 bp in length
* 52194 52293: gap of 100 bp
* 52294 65209: contig of 12916 bp in length
* 65210 65309: gap of 100 bp
* 65310 79796: contig of 14487 bp in length
* 79797 79896: gap of 100 bp
* 79897 95896: contig of 16000 bp in length
* 95897 95996: gap of 100 bp
* 95997 113419: contig of 17423 bp in length
* 113420 113519: gap of 100 bp
* 113520 133683: contig of 20164 bp in length.

FEATURES
SOURCE
1. 133683
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"

/clone="CTD-3231N5"
/clone.lib="CTD Human BAC"
1. 1459
/note="assembly-fragment"
misc_feature 1560..2520
/note="assembly-fragment"
misc_feature 2621..6334
/note="assembly-fragment"
misc_feature 6435..11913
/note="assembly-fragment"
misc_feature 12014..20756
/note="assembly-fragment"
misc_feature 20857..29735
/note="assembly-fragment"
misc_feature 29836..38645
/note="assembly-fragment"
misc_feature 38746..52193
/note="assembly-fragment"
misc_feature 52294..65209
/note="assembly-fragment"
misc_feature 65310..79796
/note="assembly-fragment"
misc_feature 79897..95896
/note="assembly-fragment"
misc_feature 95997..113419
/note="assembly-fragment"
misc_feature 113520..133683
/note="assembly-fragment"
BASE COUNT 31520 a 34761 c 35336 g 30857 t 1209 others
ORIGIN

Query Match 62.8%; Score 3142; DB 2; Length 133683;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 3609; Conservative 0; Mismatches 140; Indels 373; Gaps 7;
QY 879 ATGTCCACGACGACCAATCAGAGCTCTGTCAGATATGCAATTCGTGATTTTACT 938
DB 79796 ANNN 79855
QY 939 AGGGATGAAATTCCTCCAGACACATCTTTTCAGATAGGCGCTAGAG 998
DB 79856 NNN 79915
QY 999 AGGAGCTGAACCTACCGGGGTGACACACAGCGTGGCAAGCTGGACCAAGAACCA 1058
DB 79916 AGGAGCTGAACCTACCGGGGTGACACACAGCGTGGCAAGCTGGACCAAGAACCA 79975
QY 1059 GGACTGTGACTGACGCGGCTATTCATTTTCCATAGCCACAGGCGTCAAGAGAC 1118
DB 79976 GGACTGTGACTGACGCGGCTATTCATTTTCCATAGCCACAGGCGTCAAGAGAC 80035
QY 1119 CCAGGCGCTATTCAGAGGCTCTCTCTCTGAGAGTTCTTGGCACAGAGTTGAGGTC 1178
DB 80036 CCAGGCGCTATTCAGAGGCTCTCTCTCTGAGAGTTCTTGGCACAGAGTTGAGGTC 80095
QY 1179 AGCAGAGCCCTTACCCCACTCTCTCTCAAGGCGTGGGCTGGAACACCTGGGG 1238
DB 80096 AGCAGAGCCCTTACCCCACTCTCTCTCAAGGCGTGGGCTGGAACACCTGGGG 80155
QY 1239 AGCAGATCTTTAGCTCTGATTTTAGGGCATGATAGAGGGGCTTGGCCATAATTC 1298
DB 80156 AGCAGATCTTTAGCTCTGATTTTAGGGCATGATAGAGGGGCTTGGCCATAATTC 80215
QY 1299 CAGCCCTGGTCTGAGCCCAACCCCTCCAGAAATAATTAGAGGGGCTTGGCCAGGCTG 1358
DB 80216 CAGCCCTGGTCTGAGCCCAACCCCTCCAGAAATAATTAGAGGGGCTTGGCCAGGCTG 80275
QY 1359 TGCTAGCGTCTTCTGAGAGATTAAGAGAGGATTAAGAGAGGATCTTGTGG 1418
DB 80276 TGCTAGCGTCTTCTGAGAGATTAAGAGAGGATTAAGAGAGGATCTTGTGG 80335
QY 1419 AGGTCTGCTTAGGAGATCAATGACGGGCTCAGACTCAGCTGGGAGTCCAGCC 1478

| | | | |
|----|-------|---|-------|
| Db | 80336 | AGGTCCTGCGTTAGGGAGTCAAGTGAAGCGCGGCTCAGACATCACGTGGGCACTGCCAGCC | 80395 |
| QY | 1479 | TCCTAAGAGTGGGAGGGGCACTGGGCCACAGAGTCCAGAGAGTCCACAGCTATGTCG | 1538 |
| Db | 80396 | TCCTAAGAGTGGGAGGGGCACTGGGCCACAGAGTCCAGAGAGTCCACAGCTATGTCG | 80455 |
| QY | 1539 | CAGACCTTCTGTGGATCATCGGACCCACCTGGAAACCCACTGTGAGTACAAAGTGGCC | 1598 |
| Db | 80456 | CAGACCTTCTGTGGATCATCGGACCCACCTGGAAACCCACTGTGAGTACAAAGTGGCC | 80515 |
| QY | 1599 | CAGGTGAGTGGGCGTGGGGCTTGAAGCCCTTCAGGGTTGATGGCCATCTTGCCTATTTG | 1658 |
| Db | 80516 | CAGGTGAGTGGGCGTGGGGCTTGAAGCCCTTTCAGGGTTGATGGCCATCTTGCCTATTTG | 80575 |
| QY | 1659 | TGTGGGATATGCACACACAGGAGCATATGCGAGGTGTGTGGGACCTGTGTCTGTG | 1718 |
| Db | 80576 | TGTGGGATATGCACACACAGGAGCATATGCGAGGTGTGTGGGACCTGTGTCTGTG | 80635 |
| QY | 1719 | CAATAGCCCTAGGTTGGGGAATGAGCTTGTGTGCATCAGAGGAGCAGCAGCAGGTTG | 1778 |
| Db | 80636 | CAATAGCCCTAGGTTGGGGAATGAGCTTGTGTGCATCAGAGGAGCAGCAGCAGGTTG | 80692 |
| QY | 1779 | GCTGCAGCAAAACACACAGGGGAAAGAAAGGAGGGGCGATCAATCACTGCAAAATTTT | 1838 |
| Db | 80693 | GCTGCAGCAAAACACACAGGGGAAAGAAAGGAGGGGCGATCAATCACTGCAAAATTTT | 80738 |
| QY | 1839 | ATGAGGCTCCCGCTAAAAAAGAGGTCTCTTCTTGATGAAGAAGGAGAGAGGGG | 1898 |
| Db | 80739 | ATGAGGCTCCCGCTAAAAAAGAGGTCTCTTCTTGATGAAGAAGGAGAGAGGGG | 80738 |
| QY | 1899 | GTTTGTCTTATTAATTAAGGAGAGAGCCCGCTCAAAAAATTAAGGAGAGAGAGCC | 1958 |
| Db | 80739 | GTTTGTCTTATTAATTAAGGAGAGAGCCCGCTCAAAAAATTAAGGAGAGAGAGCC | 80738 |
| QY | 1959 | AAGACCCCGTGGTGTGTGTCTTCCAGGGGAGCTGGAACCTTTAGAGGAGCGTGG | 2018 |
| Db | 80739 | AAGACCCCGTGGTGTGTGTCTTCCAGGGGAGCTGGAACCTTTAGAGGAGCGTGG | 80738 |
| QY | 2019 | AGAACCGTGTATTCAGGCTCTCGAGAGAAAAAGAGCGGCCGCCCAAAAAATTCCTC | 2078 |
| Db | 80739 | AGAACCGTGTATTCAGGCTCTCGAGAGAAAAAGAGCGGCCGCCCAAAAAATTCCTC | 80738 |
| QY | 2079 | CCGGCGGATTAAGAAATGATGGGCTCTCTCAAAAAGATGAAGAGAAACCGAGTTGATG | 2138 |
| Db | 80739 | CCGGCGGATTAAGAAATGATGGGCTCTCTCAAAAAGATGAAGAGAAACCGAGTTGATG | 80738 |
| QY | 2139 | TGTTGATATTTTAAAACTCAGGTAGNNNNNNNNNTGCTTCAGTAAATTTTATTTAG | 2198 |
| Db | 80739 | TGTTGATATTTTAAAACTCAGGTAGNNNNNNNNNTGCTTCAGTAAATTTTATTTAG | 80761 |
| QY | 2199 | CGCCTTCTAGAGAAACAAAGAGGCTTCCATTTCTGAGAGAGAAACAGCAGAGAAACAG | 2258 |
| Db | 80762 | CGCCTTCTAGAGAAACAAAGAGGCTTCCATTTCTGAGAGAGAAACAGCAGAGAAACAG | 80821 |
| QY | 2259 | GCAGATATCCTGTATTAATTTCAAGATGATTAAGGCTCTCTAGAAATATCAAGCAAGGT | 2318 |
| Db | 80822 | GCAGATATCCTGTATTAATTTCAAGATGATTAAGGCTCTCTAGAAATATCAAGCAAGGT | 80881 |
| QY | 2319 | GAGAGACAGAGACACCGGCTGGCAGTGGGCTATTTCCAGGTTGATGGTTGGAGC | 2378 |
| Db | 80882 | GAGAGACAGAGACACCGGCTGGCAGTGGGCTATTTCCAGGTTGATGGTTGGAGC | 80941 |
| QY | 2379 | ATCCTTTCTTAAAGGAACCTGAGTGGGAAAGGAACATGACGATATCTCAGAGAGGCT | 2438 |
| Db | 80942 | ATCCTTTCTTAAAGGAACCTGAGTGGGAAAGGAACATGACGATATCTCAGAGAGGCT | 81001 |
| QY | 2439 | CTTCCAGGAGAAATCAGACAGTGGAAAGGCCCTGAGGACACATTCAGTAAACATCA | 2498 |
| Db | 81002 | CTTCCAGGAGAAATCAGACAGTGGAAAGGCCCTGAGGACACATTCAGTAAACATCA | 81061 |
| QY | 2499 | TTTGAGCATCTTACACGCTAGTTCATTTATGGGAATGGGAATATGTTGGTGGAGGG | 2558 |
| Db | 81062 | TTTGAGCATCTTACACGCTAGTTCATTTATGGGAATGGGAATATGTTGGTGGAGGG | 81121 |

[illegible]


```

* 1927 2026: gap of unknown length
* 2027 4202: contig of 2176 bp in length
* 4203 4302: gap of unknown length
* 4303 7163: contig of 2861 bp in length
* 7164 7263: gap of unknown length
* 7264 9961: contig of 2698 bp in length
* 9962 10061: gap of unknown length
* 10062 12576: contig of 2514 bp in length
* 12576 12676: gap of unknown length
* 12676 15545: contig of 2870 bp in length
* 15545 15646: gap of unknown length
* 15646 20942: contig of 5296 bp in length
* 20942 21042 23966: contig of 2925 bp in length
* 23967 24066: gap of unknown length
* 24067 27231: contig of 3165 bp in length
* 27232 27332 30550: contig of 3219 bp in length
* 30551 30651 35202: gap of unknown length
* 35203 35303 40201: contig of 4552 bp in length
* 40202 40302 45758: gap of unknown length
* 45759 45859 51249: contig of 5457 bp in length
* 51250 51349: gap of unknown length
* 51350 58267: contig of 5391 bp in length
* 58268 58367: gap of unknown length
* 58368 66916: gap of 6918 bp in length
* 66917 67016: contig of 8549 bp in length
* 67017 77145: gap of unknown length
* 77146 77245: contig of 10129 bp in length
* 77246 86493: gap of unknown length
* 86494 86594 96688: gap of 9248 bp in length
* 86594 96689 96789: gap of unknown length
* 96789 107169: gap of 10095 bp in length
* 107170 107270 107269: gap of unknown length
* 107271 117047: contig of 10381 bp in length
* 117048 117147: gap of unknown length
* 117148 127120: contig of 9778 bp in length
* 127121 127220: gap of unknown length
* 127221 142029: gap of 9973 bp in length
* 142030 142128: contig of unknown length
* 142129 160169: gap of 14809 bp in length
* 160169 160169: contig of 18040 bp in length.
FEATURES
Location/Qualifiers
1. 160169
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-810P12"
1. 1926
/feature="assembly_name:Contig18"
2027. 4202
/feature="assembly_name:Contig19"
4303. 7163
/feature="assembly_name:Contig20"
clone_end:77
vector_side:left"
7264. 9961
/feature="assembly_name:Contig21"
10062. 12575
/feature="assembly_name:Contig22"
12676. 15545
/feature="assembly_name:Contig23"
15646. 20941
/feature="assembly_name:Contig24"
21042. 23966
/feature="assembly_name:Contig25"
24067. 27231
/feature="assembly_name:Contig26"
27332. 30550

```

```

/feature="assembly_name:Contig27"
30651. 35202
/feature="assembly_name:Contig28"
35303. 40201
/feature="assembly_name:Contig29"
40302. 45758
/feature="assembly_name:Contig30"
45859. 51249
/feature="assembly_name:Contig31"
51350. 58267
/feature="assembly_name:Contig32"
58368. 66916
/feature="assembly_name:Contig33"
67017. 77145
/feature="assembly_name:Contig34"
77246. 86493
/feature="assembly_name:Contig35"
86594. 96688
/feature="assembly_name:Contig36"
96789. 107169
/feature="assembly_name:Contig37"
107270. 117047
/feature="assembly_name:Contig38"
117148. 127120
/feature="assembly_name:Contig39"
vector_end:SP6
/feature="assembly_name:Contig40"
142130. 160169
/feature="assembly_name:Contig41"
160169. 160169
BASE COUNT 41647 a 36693 c 37296 g 42226 t 2307 others
ORIGIN
Query Match 47.6% Score 2379.8; DB 2; Length 160169;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 2840; Conservative 0; Mismatches 122; Indels 367; Gaps 8;
QY 817 CACTCTCTGTAAGCACTACTGATGAGGCCCTGCAGCAATCAGCCAGAAATACGTA 876
DB 4199 CAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4258
QY 877 TGATGTCACACAGCAATCAGAGCTCTCTGATGCAATGCAATGTCATTTTA 936
DB 4259 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4318
QY 937 CTAGGCTGATGAATTCACACACACATCTCTTTCAGATPAAGGCACTGAGCTGAG 996
DB 4319 CTAGGCTGATGAATTCACACACACATCTCTTTCAGATPAAGGCACTGAGCTGAG 996
QY 997 AGAGAGCTGAACCTACCCGGGTCACACACAGGAGGCGAAGGCTGGACAGAAAC 4378
DB 4379 AGAGAGCTGAACCTACCCGGGTCACACACAGGAGGCGAAGGCTGGACAGAAAC 4378
QY 1057 CAGAGCTGTAAGTACCTGACCCGGATTCATCTTTCATAGCCAGGCGTCAAGA 1116
DB 4439 CAGAGCTGTAAGTACCTGACCCGGATTCATCTTTCATAGCCAGGCGTCAAGA 1116
QY 1117 CCCCAGGCGCTAGTCAGAGGCTCTCTCTCTGAGAGGCTCTGAGAGGCTCTGAGAGG 4497
DB 4498 CCCCAGGCGCTAGTCAGAGGCTCTCTCTCTGAGAGGCTCTGAGAGGCTCTGAGAGG 4497
QY 1177 TCAGCAAGCCCCCTAACCCCACTCTCTGCAAGGCGTCAAGAGGCTCTGAGAGG 1236
DB 4558 TCAGCAAGCCCCCTAACCCCACTCTCTGCAAGGCGTCAAGAGGCTCTGAGAGG 1236
QY 1237 GGAGCAGATCTTATGAGCTCTGATTTTGAAGGCGTCAAGAGGCGTCTGAGAGG 1296
DB 4618 GGAGCAGATCTTATGAGCTCTGATTTTGAAGGCGTCAAGAGGCGTCTGAGAGG 1296
QY 1297 TCAGGCGCTGCTCAGCCCAACACCCCTCAAGAAATTAAGAGGCGCATGAGCC 1356
DB 4678 TCAGGCGCTGCTCAGCCCAACACCCCTCAAGAAATTAAGAGGCGCATGAGCC 1356

```

| | | | |
|----|------|--|------|
| OY | 1357 | TGTCGTAGCGCGTTCCTCTGTAGCAGATTACAAAGAGGAGCTTAAGACAAAGGACCTCTTGT | 1416 |
| Db | 4737 | TGTCGTAGCGCGTTCCTCTGTAGCAGATTACAAAGAGGAGCTTAAGACAAAGGAGCTCTTGT | 4796 |
| OY | 1417 | GGAAGCTCTGGCTTAGGGAATCAAGTAGCGCGGCTCAGCACTCAGCTGGGAGTGGCAG | 1476 |
| Db | 4797 | GGAGGTCTGGCTTAGGGAATCAAGTAGCGCGGCTCAGCACTCAGCTGGGAGTGGCAG | 4856 |
| OY | 1477 | CCCTTAGAGTGGGCAGAGGGGCACCTGGGCCAAGAGATCCCAAGGAGTCCCAACAGCTAGTC | 1536 |
| Db | 4857 | CCCTTAGAGTGGGCAGAGGGGCACCTGGGCCAAGAGATCCCAAGGAGTCCCAACAGCTAGTC | 4916 |
| OY | 1537 | GCCAGACCTTCTGTGGGATCATGGAGCCACCTTGGAAACCCACCTGTAGTACAAAGGTGC | 1596 |
| Db | 4917 | GCCAGACCTTCTGTGGGATCATGGAGCCACCAOCTTGGAAACCCCAOCTGTAGTACAAAGGTGC | 4976 |
| OY | 1597 | CCCAAGTGGACATGGGCTTGGGGCTTTAGAGCCTTCAGGGTTGGATGGCCATCTTGCATATT | 1656 |
| Db | 4977 | CCCAAGTGGACATGGGCTTGGGGCTTTAGAGCCTTCAGGGTTGGATGGCCATCTTGCATATT | 5036 |
| OY | 1657 | TGTCGTGGGAATGCACACACAGGCGACACATGCCACAGTGTGTGGGCACTGTGTCTGT | 1716 |
| Db | 5037 | TGTCGTGGGAATGCACACACAGGCGACACATGCCCAAGTGTGTGGGCACTGTGTCTGT | 5096 |
| OY | 1717 | TGCAAATGCCCTGAGGTGGGAATGAGCTTGGTGTGCATCAAGAGCGACAGCCAGCCAGTG | 1776 |
| Db | 5097 | TGCAAATGCCCTGAGGTGGGAATGAGCTTGGTGTGCATCAAGAGCGACAGCCAGCCAGTG | 5153 |
| OY | 1777 | TGCGTCGAGCAAAAACACACAGGCAAAATGAGAGGGGCAATCAATCACTGCACAAAATATT | 1836 |
| Db | 5154 | TGCGTCGAGCAAAAACACACAGGCAAAATGAGAGGGGCAATCAATCACTGCACAAAATATT | 5201 |
| OY | 1837 | TTATAGAGCTCCCCCTAAAAAAGAAAGAGTCTTTTGGATGAAAGAGGAGAGAGG | 1896 |
| Db | 5202 | ----- | 5201 |
| OY | 1897 | GGGTTTGCCTTAATAATATAGAGGAGAGCGCCCTCAAAAAATAGAGGAGGAGAGC | 1956 |
| Db | 5202 | ----- | 5201 |
| OY | 1957 | CCAAGACCCCGTGGGTTGTGTTCAGGGGAGCTCGAACCTTTAGAGGAGCGTG | 2016 |
| Db | 5202 | ----- | 5201 |
| OY | 2017 | GGAGAACCGGTGATTTCAGGCTCTCGAGAGAAAGAGGGCGCCCAAAAATATATCC | 2076 |
| Db | 5202 | ----- | 5201 |
| OY | 2077 | TCCGGGCGATTAAGAAATGGTGCCCTCTCTCAAAAAGTGAAGAGAAAGCCGGAATTGTA | 2136 |
| Db | 5202 | ----- | 5201 |
| OY | 2137 | TGTGTTGATATTTTAAAACTCCAGGTAGNNNNNNNNNGCTTCAGTAAATTTTATGTG | 2196 |
| Db | 5202 | -----TCTTCAGTAAATTTTATGTG | 5222 |
| OY | 2197 | AGCGCCCTTCTACGAAACACAAAGAGAGCTTCATTTGAGAGAGAAACAGCGCAGAAAC | 2256 |
| Db | 5223 | AGCGCCCTTCTACGAAACACAAAGAGAGCTTCATTTGAGAGAGAAACAGCGCAGAAAC | 5282 |
| OY | 2257 | AGGAGATATTCCTTATATTTCAAGTAGTAGTAAAGTCTCTTAGAAATTTCAAGCAAG | 2316 |
| Db | 5283 | AGGAGATATTCCTTATATTTCAAGTAGTAGTAAAGTCTCTTAGAAATTTCAAGCAAG | 5342 |
| OY | 2317 | GTGAGGAGACACAGAGCAGCGGTGGAGTGGGGCTCATTTCCAGGTTGAGTGGTGGGA | 2376 |
| Db | 5343 | GTGAGGAGACACAGAGCAGCGGTGGAGTGGGGCTCATTTCCAGGTTGAGTGGTGGGA | 5402 |
| OY | 2377 | ACATCTTTTCTAAAGGAACTGCGATGGGAAAGAACCATTCAGGATCTCAGAGAAAGC | 2436 |
| Db | 5403 | ACATCTTTTCTAAAGGAACTGCGATGGGAAAGAACCATTCAGGATCTCAGAGAAAGC | 5462 |

| | | | |
|----|------|--|------|
| OY | 2437 | TTCCTCCAGGCACGAAGATCAGAGGTGGAAAAGGCCCTGAGACCACCAATTGACTGAATAAT | 2496 |
| Dd | 5463 | TTCTCTCCAGGCAGAAGATCGAGAGGTGGAAAAGGCCCTTGAGACCACCAATTGACTGAATAAT | 5522 |
| OY | 2497 | CATTGTGAGCATCTCTACACGCTAAGTTCATTATATGGAAATGGGAAATATGCTGTGGACAG | 2556 |
| Dd | 5523 | CATTGTGAGCATCTCTACACGCTAAGTTCATTATGGAATGGGAATATGCTGTGGACAG | 5582 |
| OY | 2557 | GCGTGCCTGGTCCCCTTCATCTCTCTCACACTAAGGTGGTTTAGAGACTCTGGAGCTCAA | 2616 |
| Dd | 5583 | GCGTGCCTGGTCCCCTTCATCTCTCTCACACTAAGGTGGTTTAGAGAGCTGTGGAGCTPAA | 5642 |
| OY | 2617 | CGAACAAGATGGGCTGAGAACACTGGCTGAGCCAGAGACGTGACCTTAAGTGTAGACA | 2676 |
| Dd | 5643 | CGAACAAGATGGGCTGAGAACACTGGCTGAGCCAGAGACGTGACCTTAAGTGTAGACA | 5702 |
| OY | 2677 | TTGCTGCTGTACTGCTCTTGTCTGTGTATATTTATTTATTTATTTATTTATGATCTTAAGA | 2736 |
| Dd | 5703 | TTGCTGCTGTACTGCTCTTGTCTGTGTATATTTATTTATTTATTTATTTATTTATTTATTTGA | 5762 |
| OY | 2737 | CAGAGTTTTGCTCTTCCTTCAACCCAGGGTTTAGAGCAATGSGCGATCTCAGCTCAGCTGCAA | 2796 |
| Dd | 5763 | CAGAGTTTTGCTCTTCTTTACACAGGCTGGAGTGCATAATGSGCGATCTCAGCTCAGCTGCAA | 5822 |
| OY | 2797 | CCTCCACCTCTCTGGGATCAAGGAGTTCTCTGCTCAGCTCCTCGTAGTACGTGGGATTAC | 2856 |
| Dd | 5823 | CCTCCACCTCTCTGGGTTCAAGGCAATTCCTCGCTCAGGCTCCTGATGCGGGGATTAC | 5882 |
| OY | 2857 | AGGACACC CGGACACAGCGCTGGATAATTTTTTTGTATTTTATAGTAGAGAGGGTTTACCC | 2916 |
| Dd | 5883 | AGGACCCCCGACACAGCGCTGAGATAATTTTTTTGTATTTTATAGTAGAGAGGGTTTACCC | 5942 |
| OY | 2917 | ATGTTGGCCAGCGCTGCTCTGCAATCTCTGACTTAGGTATCCACTGCTTGACTTCCC | 2976 |
| Dd | 5943 | ATGTTGGCCAGCGCTGGTCTCTGCAATCTCTGACTTAGGTATCCACTGCTTGACTTCCC | 6002 |
| OY | 2977 | AAAGTGTGGGATTTATAGGATGAGCAGTGGGCCAGAGTATATAGAAAGTTAAAGCA | 3036 |
| Dd | 6003 | AAAGTGTGGGATTTATAGGATGAGCAGTGGGCCAGAGTATATAGAAAGTTAAAGCA | 6062 |
| OY | 3037 | CATGGCAATGCACAGCGCTATCTAGCTCTTCCCTGGCAAGCAAAAGGCGCTCTTG6GC | 3096 |
| Dd | 6063 | CATGGCAATGCACAGCGCTATCTAGCTCTTCCCTGGCAAAAGGCGCGCTCTTG6GC | 6122 |
| OY | 3097 | TCACCTTCTTGCGTTCTCTACTTCCAAAAAGGCAATGCAATGSGAGGCGCTTGAGACCA | 3156 |
| Dd | 6123 | TCACCTTCTTGCGTTCTCTACTTCCAAAAAGGCAATGCAATGSGAGGCGCTTGAGACCA | 6182 |
| OY | 3157 | CTTCATCCACCTCTAGAGGTCCCTATAGGGAGAGTTAGGTCGAGACAGGAGAGGGTCT | 3216 |
| Dd | 6183 | CTTCATCCACCTCTAGAGGTCCCTATAGGGAGAGTTAGGTCGAGACAGGAGAGGGTCT | 6242 |
| OY | 3217 | GACAGGCTCTGACAGGGCTCTGATTCCTACAAACCCCAATCGGTGTCCCTCTCTAAC | 3276 |
| Dd | 6243 | GACAGGCTCTGACAGGGCTCTGATTCCTACAAACCCCAATCGGTGTCCCTCTCTAAC | 6302 |
| OY | 3277 | AGGACCCCAAGCCCAAGCGCTGAGCCGACACTGCGTGGSCATATGACATCACTTACAAAGCC | 3336 |
| Dd | 6303 | AGGACCCCAAGCCCAAGCGCTGAGCCGACACTGCGTGGSCATATGACATCACTTACAAAGCC | 6362 |
| OY | 3337 | AAGTGGCTAAATGCCCGTTAAGGCTCTCTCCCGCTGCTGCTGTGCTGTGGGGGGACGA | 3396 |
| Dd | 6363 | AAGTGGCTAAATGCCCGTTAAGGCTCTCTCCCGCTGCTGCTGTGCTGTGGGGGGACGA | 6422 |
| OY | 3397 | TCCTCAAGGCGATATAGGCGAGTTCCTATCTCTGCTGCGCTCTCAATCAATCCGT | 3456 |
| Dd | 6423 | TCCTCAAGGCGATATAGGCGAGTTCCTATCTCTGCTGCTCTCAATCAATCCGT | 6482 |
| OY | 3457 | TTATTTTATAGTAAAGCTGGCAGGGCTGGGCCGGGGGACCTGGGAAGATGTGCTGGGG | 3516 |
| Dd | 6483 | TTATTTTATAGTAAAGCTGGCAGGGCTGGGCCGGGGGACCTGGGAAGATGTGCTGGGG- | 6541 |
| OY | 3517 | CTGGGAGCTGGAGCTCTTG6GGGCCCTCCAGCAGCACTCAGGGGCCAGTGCACACAGTCCA | 3576 |

[illegible]

| | | | |
|----|------|---|------|
| QY | 2042 | CGAGAGAAAGAGGAGCGCCGCAAAAATAATCCCTCCCGGGCCATAGAAATGCTGCC | 2101 |
| Db | 5292 | ----- | 5291 |
| QY | 2102 | TCTCTCAAAAAGATGAAGAGGAGCCGGAGTTGATGTGATATTTTAAACTCCAG | 2161 |
| Db | 5292 | ----- | 5291 |
| QY | 2162 | GTACNNNNNNNNNNNGCTCTCAGTAAATTTTATTAGCGCCTTCTACGAGAACACAG | 2221 |
| Db | 5292 | -----TTGTTTATAGTAAATTTTATTGAGCGTTTATTACAGATATTAAGAG | 5338 |
| QY | 2222 | GAGCTCCATCTGAGAGAGAAACAGCAGAGAAACAGCAGATATCCGTATATTTCA | 2281 |
| Db | 5339 | GAGTTTTTATTAGAGAGAAATAGTAGAAATAGTAAATATTTTGTAAATTTTAA | 5398 |
| QY | 2282 | GTAAGTATAGTCTCTCTAGAAATATCAGCAAGGTGAGAGACACAGACCAGGTG | 2341 |
| Db | 5399 | GTAATGTAACTGTTTTTAAATAATTAATAGTAGTAGAGATATAGATATCGGTG | 5458 |
| QY | 2342 | CAGTGGCGCTATATTTCCAGGTGTGATGGTGGGAACATCCCTTAAGGAACCTGGA | 2401 |
| Db | 5459 | TAGTGGGTTTATTATTAGTTAGTTGATGTTGGAAATATTTTAAAGGAATTTGGA | 5518 |
| QY | 2402 | GTGGGAAGAACATGACAGATATCTCAGGAAGACCTTCCACAGCAAGAAATACAG | 2461 |
| Db | 5519 | GTGGGAAGAAATATATGTAGTATTTTGAAGAGACTTTTTTTAGTAGGAATATAGAG | 5578 |
| QY | 2462 | GTGGAAGGCCCTGAGCCACCACATTCAGTAAACATATTAGATCTTACCCAGTACG | 2521 |
| Db | 5579 | GTGGAAGGTTTTTGGAGTATATTATTTAGTAATATTTATGAGATATTTATTTAGTACG | 5638 |
| QY | 2522 | TTCCATTTATGGGAATGGGAATATGTTGGTGACAGGCTGCTGCTCCATCTCATCTC | 2581 |
| Db | 5639 | TTTTATTTATGGGAATGGGAATATGTTGGTGATAGGTTGTTTGTTTTTTATATTTT | 5698 |
| QY | 2582 | TCACACTAGGTTGGTTGAGAGAGCTTGGGAGCTAACGAAACAGATGGCTGAGAACTG | 2641 |
| Db | 5699 | TTAATATTAGGTTGGTTGAGAGAGTTGGGAATTACGAAATAGATGGGTTGAGAAATTTG | 5758 |
| QY | 2642 | CTTACCCAGAGAGACTGAGCTTATGTGTGATATGCTGCTCTTACCTGCTTGGCT | 2701 |
| Db | 5759 | TTTATGTTTAGAGATTTGACAGTTGATGTGTGATATGTTGTTTATTTTGTATTTT | 5818 |
| QY | 2702 | TGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAGG | 2761 |
| Db | 5819 | TGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAGG | 5878 |
| QY | 2762 | CTTAGAGCAATGGCGTATCTCAGCTCACGCAACTCCACCTCCGGAATCAAGGAT | 2821 |
| Db | 5879 | TTGAGAGTAAATGGCGTATTTAGTTATTTGTAATTTATTTTGGGTTAAAGGAT | 5938 |
| QY | 2822 | TCTCCCTCCTCAGCGCTCCTGAGTACGTGGGATTAACAGGACCCGACACCGCTGATAA | 2881 |
| Db | 5939 | TTTTTTCTTTAGTTTTTTGAGTGGGATTTATAGTAATCTATTTAGCTTTGGATTA | 5998 |
| QY | 2882 | TTTTTTGTATTTTATAGACAGGCTTACACATGTTGGCCAGGCTGCTCGAACT | 2941 |
| Db | 5999 | TTTTTTGTATTTTATAGAGAAAGGTTTTATTAATGCTTGGTTATGTTTCCGAAAT | 6058 |
| QY | 2942 | CTGACACTTATAGTATCCACTGCTGCGACTTCCCAAGTGTGGGATTTATAGGCAAG | 3001 |
| Db | 6059 | TTTATTTATTTAGTATTTATTTGTTGATTTTTTAAAGTGTGGGATTTATAGGTATAG | 6118 |
| QY | 3002 | CCACTGGCCAGTATATAGAAAGTTAAAGCACATGGCAATGCAACAGCCTATAC | 3061 |
| Db | 6119 | TTATTTGCGTTTATGATTTATAGAAAGTTAAAGTATATGTAATGTATATTTATAC | 6178 |
| QY | 3062 | GTCTTCCCTCAGAAAGAGGAGCGCTCTGGGCTACATTTCTTGGCTTTCTACTTCCA | 3121 |
| Db | 6179 | GTATTTTATTTTAAAGGAGGATTTTGGGTTATTTTATTTTGGCTTTATTTATTTTAA | 6238 |

| | | | | |
|------------|--|--------|---|------|
| Oy | | 3122 | AAAGCGACGTGACATGGCAGGCGCTTGGGAACCACTTCACCTCCCTGCCTTA | 3181 |
| Dd | | 6239 | AAAGCTGATTGAAATTTGGTAGGGCTTTTGGAAGATTATTTATTTTAAAGGCTTTT | 6298 |
| Oy | | 3182 | TGGGAGAGTTGAGTCCAGAACGAGGAGGCTTCGACAGGCTCTGACAGAGGCTCTGA | 3241 |
| Dd | | 6299 | TGGGAGAGTTGAGTTCAGATTAGTAGGAAAGGCTTTTGAATAGGCTTTTGAATAGGCTTTTGA | 6358 |
| Oy | | 3242 | TCCTACAAAACCCCATCGGTCTCCCTCTCTACACAG | 3279 |
| Dd | | 6359 | TTTTTATAAATTTTAAATCGGTCTTTTAAAAATTAAG | 6396 |
| RESULT 10 | | | | |
| LOCUS | AC051664/c | | | |
| DEFINITION | Homo sapiens chromosome 11 clone RP11-810P12, WORKING DRAFT | | | |
| ACCESSION | AC051664 | | | |
| VERSION | AC051664.2 | | | |
| KEYWORDS | HTG; HTGS_PHASE1; HTGS_DRAFT. | | | |
| SOURCE | Homo sapiens. | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| REFERENCE | Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. | | | |
| AUTHORS | 1 (bases 1 to 160169) | | | |
| TITLE | Waterston,R.H. | | | |
| REFERENCES | The sequence of Homo sapiens clone | | | |
| JOURNAL | Unpublished | | | |
| REFERENCE | 2 (bases 1 to 160169) | | | |
| AUTHORS | Waterston,R.H. | | | |
| TITLE | Direct Submission | | | |
| JOURNAL | Submitted (15-APR-2000) Genome Sequencing Center, Washington | | | |
| COMMENT | University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA On Apr 23, 2000 this sequence version replaced gl:7574984. | | | |
| | ----- Genome Center ----- | | | |
| | Center: Washington University Genome Sequencing Center | | | |
| | Center code: WUGSC | | | |
| | Web site:http://genome.wustl.edu/gsc/index.shtml | | | |
| | Project Information ----- | | | |
| | Center project name: H.NH0810P12 | | | |
| | Summary Statistics ----- | | | |
| | Sequencing vector: M13; 100% | | | |
| | Sequencing vector: plasmid; 0% | | | |
| | Chemistry: Dye-primer ET; 100% of reads | | | |
| | Chemistry: Dye-terminator Big Dye; 0% of reads | | | |
| | Assembly program: Phrap; version 0.990319 | | | |
| | Consensus quality: 147316 bases at least Q40 | | | |
| | Consensus quality: 151042 bases at least Q30 | | | |
| | Consensus quality: 152875 bases at least Q20 | | | |
| | Insert size: 165000; agarose-fp | | | |
| | Insert size: 157869; sum-of-contigs | | | |
| | Quality coverage: 3.91 in Q20 bases; agarose-fp | | | |
| | Quality coverage: 4.17 in Q20 bases; sum-of-contigs | | | |
| | ----- NOTE: This is a 'working draft' sequence. It currently | | | |
| | * consists of 24 contigs. The true order of the pieces | | | |
| | * is not known and their order in this sequence record is | | | |
| | * arbitrary. Gaps between the contigs are represented as | | | |
| | * runs of N, but the exact sizes of the gaps are unknown. | | | |
| | * This record will be updated with the finished sequence | | | |
| | * as soon as it is available and the accession number will | | | |
| | * be preserved. | | | |
| * | 1 | 1926: | contig of 1926 bp in length | |
| * | 1927 | 2026: | gap of unknown length | |
| * | 2027 | 4202: | contig of 2176 bp in length | |
| * | 4203 | 4302: | gap of unknown length | |
| * | 4303 | 7163: | contig of 2861 bp in length | |
| * | 7164 | 7263: | gap of unknown length | |
| * | 7264 | 9961: | contig of 2698 bp in length | |
| * | 9962 | 10061: | gap of unknown length | |

Db 76647 TTGTAATAATATACAAATGAAATCTTGTCTTAAATATGCTAGCAACAGCACACA 76588
Qy 4764 TTTGCTTTTACTAGGACACCAAGCTACTTTAAAAAGTTAGGCGACCAAGGGCT 4823
Db 76587 TTTGGTTTTTACTAGGGACCAAGGACTTTAAAAAGTTAGGCGACCAAGGGCT 76528
Qy 4824 CACACCTGTATATCCCGACACTTTGGAGGCGCAAGAGAGATCATCTTGGCCAGAG 4883
Db 76527 CACACCTGTATATCCCGACACTTTGGAGGCGCAAGAGAGATCATCTTGGCCAGAG 76468
Qy 4884 TTTAGACCTGTAGCAACATAGGAGATCTGTCTATATAAAAATTTAAAAAT 4943
Db 76467 TTTAGACCTGTAGCAACATAGGAGATCTGTCTATATAAAAATTTAAAAAT 76408
Qy 4944 GCGTAGGCGCTTTGGCTTACACCCGTAATCCCGACACTTTGGAGGCGCGGGGT 5000
Db 76407 G--CTAGGCCCTTGTCTACACCCGTAATCCCGACACTTTGGAGGCGCGGGGT 76353

RESULT 11
AC084857/c 133683 bp DNA linear HTG 22-NOV-2000
LOCUS Homo sapiens chromosome 11 clone CTD-3231N5 map 11, WORKING DRAFT
DEFINITION
SEQUENCE, 13 unordered pieces.
AC084857
AC084857.1 GI:11276215
HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
Homo sapiens.
SOURCE
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
1 (bases 1 to 133683)
Mammalia: Eutheria; Primates; Catarrhini; Homnidae; Homo.
Birren, B., Linton, L., Nussbaum, C. and Lander, E.
Homo sapiens chromosome 11, clone CTD-3231N5
Unpublished
2 (bases 1 to 133683)
Birren, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bede, F., Boguslavsky, L.,
Boukhalter, B., Brown, A., Burkett, G., Campolano, A., Castle, A.,
Choedel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Deavelano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Glade, S., Goyette, M.,
Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Labocque, K.,
Lamazares, R., Landers, T., Lechocsky, J., Levine, R., Lieu, C., Liu, G.,
Macdonald, P., Margolis, N., McCarthy, M., McEwan, P., McKernan, K.,
McPheters, R., Meldrum, J., Meneses, L., Mihova, T., Mlenga, V.,
Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
O'Donnell, P., O'Neill, D., Olyar, T.M., Oliver, J., Peterson, K.,
Pierre, N., Pisanl, C., Pollara, V., Raymond, C., Ribback, M., Riley, R.,
Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
Sougniez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Sudramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Tirrell, A., Travers, M., Trifillo, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (22-NOV-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02144, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: W1BR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: l11038
Center clone name: 3231_N5
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads

FEATURES

source

misc_feature
1..1459
/note="assembly_fragment"
misc_feature
1560..2520
/note="assembly_fragment"
misc_feature
2621..6334
/note="assembly_fragment"
misc_feature
6335..11913
/note="assembly_fragment"
misc_feature
11914..12013
/note="assembly_fragment"
misc_feature
12014..20756
/note="assembly_fragment"
misc_feature
20757..29735
/note="assembly_fragment"
misc_feature
29736..29835
/note="assembly_fragment"
misc_feature
29836..38645
/note="assembly_fragment"
misc_feature
38646..52193
/note="assembly_fragment"
misc_feature
52194..52293
/note="assembly_fragment"
misc_feature
52294..65209
/note="assembly_fragment"
misc_feature
65210..65309
/note="assembly_fragment"
misc_feature
65310..79796
/note="assembly_fragment"
misc_feature
79797..79897
/note="assembly_fragment"
misc_feature
79898..95996
/note="assembly_fragment"
misc_feature
95997..113419
/note="assembly_fragment"
misc_feature
113420..113520
/note="assembly_fragment"

Identified using the e-PCR software (G. Schuler)"

| | | | |
|----|-------|---|-------|
| OY | 4605 | CTCAAAAGCAAAACAAATTAATGTTTAATCTAAACGGTAAGGTAAATTAATTCAGAAATATATGA | 4664 |
| Db | 71002 | GAGAGATCTCTGTCTC(A)AAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAT | 71061 |
| OY | 4665 | TAGCATTTAAATTGAAAAAGCATTAAT-----GATTCATAGATTTGTAAAA | 4711 |
| Db | 71062 | CACCGAAGCAACCAATACAGAACTCATCTGGGCTGAAGAGGCCAATCAGAACTGAGCTA | 71122 |
| OY | 4712 | TATCAAAATACATGAATATCTTGTTGTTTAATAATGTCAGCAACAAAGCATTTGATT | 4771 |
| Db | 71122 | TATTGACCAATCAGAACTAAGCTAGTTTGAGCGCTTCATTTGCATTAATGAGACCTATGG | 71181 |
| OY | 4772 | TTACTAGGCGACCAAG--GTACTTTAAAAAATTAGGCGCACCCACAG-----GGCTC | 4824 |
| Db | 71182 | GAACTTGGGCGAGAACTTGTGCTATATAAAACCCGAACTGTCCACCTGAGCGGATGGCTTC | 71241 |
| OY | 4825 | ACACCTCTAATCCACGACTTTGGGAGGCCA-----GGCAGGAGATCAGCTTGAGGCC | 4878 |
| Db | 71242 | ACAATCTCTAATCCACGACTTTGGGAGGCCAAGGGTGGGGGGGCGAGATCACTGAGGTC | 71301 |
| OY | 4879 | AGAGATTAGGACTGTAGCAACAATAGGAGATCTCTATCTCTCTATATAAAATTAATA | 4938 |
| Db | 71302 | AGAGATTGCGAGACACGCTGACCAACATGG---TGAATCTCTCTACTATAAAATATACAA | 71358 |
| OY | 4939 | AAATTGGCTAGGCCCTTTGGCTTACACACCGGTATCCACACACTTTGGAGGCCGAGGC | 4996 |
| Db | 71359 | AAATTGACGAGACCTGTGGTGTGTCACCTGTATGCTCCAGATACTCTGGGAGGGCTGAGGC | 71416 |

[illegible]

| | | | | |
|----|--|-------|--|-------|
| Dd | | 33585 | TGAGAGAGATCCTGTCTCAAAAACAACAAACAACAAACAACAAACAACAAACAAC | 3364 |
| Oy | | 4662 | TGATAGCATTTTAAATTGAAAAAACCATTAAT-----GATTACATGGATTGTA | 4708 |
| Dd | | 33645 | ATTCCACTGAACAACAACCAATCAAGAATCAATCTGGGTGAAGAGGCCAATTCAGAACTGAG | 3370 |
| Oy | | 4709 | AAATATCAAAATACATGAAATTCCTGTGTCTTAATTAATGCATAGCAACAAGCACATTGG | 4768 |
| Dd | | 33705 | CATATTTAGACCATAACAGAACTAAGCTTAGCGCTTCATTGGCATTAATGGACCTGA | 3376 |
| Oy | | 4769 | TTTTTACTAGGGCACCAAG--GTACTTTAAAAAAAGTTAGGGCCAGCCACAGG---- | 4821 |
| Dd | | 33765 | TTGGAACCTGGGCGAGAACTGTGCTATTAAGGCCAAGATCTGCAGCTGGACGGGATGG | 3382 |
| Oy | | 4822 | CTCACCACCTGTAAATCCAGACACTTTGGGAGGCCA-----GGCAGAGCATCACTTGAG | 4875 |
| Dd | | 33825 | CTCACATCTGTAAATCCAGACACTTTGGGAGGCCAAGGGGTGGGGGGCCATATCATCTGAG | 3388 |
| Oy | | 4876 | CCCAGGAGTTTAGGACCTGAGACCAATAGGAGAGATCCTGATCTTGCTTATTAATAATT | 4935 |
| Dd | | 33885 | GTCAGGAGTTCGAGACCTGACCTGACCAACATGG---TGAAACTGTCTTACTTAATAATA | 3394 |
| Oy | | 4936 | AAAAAATGGCTAGGCCCCCTTTGGCTTACACCCCGTAACTCCAGACATTGGGAGCCGAGG | 4995 |
| Dd | | 33942 | CAAAAATTAGACGAGCACTGTGGTGGTGCACCTGTAGTCCAGATACTGGGAGGCTGAGG | 3400 |
| Oy | | 4996 | C | 4996 |
| Dd | | 34002 | C | 34002 |

RESULT 14
HSJ309K20/c

| | | | | | |
|------------|---|----------|------------|--------|-----------------|
| LOCUS | HSJ309K20 | 96599 bp | DNA | linear | PRI 02-AUG-2001 |
| DEFINITION | Human DNA sequence from clone RP1-309K20 on chromosome 20 Contains the gene for a novel protein similar to dyfsterlin, the SPAG4 gene for sperm associated antigen 4, the CPHEI gene for Cope I (similar to KIAA0636), the gene KIAA0765 (HRHHPF2091) for an RNA recognition motif (RNP, RRM or RBD domain) containing protein and the 3' end of the NIFS gene for cysteine desulfurase. Contains ESTs, STSs, GSSs and four putative CpG islands, complete sequence. | | | | |
| VERSION | AL109827 | 8 | GI:6706902 | | |
| KEYWORDS | HG: AGR5; cogene; Cpg island; cysteine desulfurase; dyfsterlin; KIAA0765; NIFS; RBD; RNP; RRM; SPAG4. | | | | |
| SOURCE | human. | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | |
| AUTHORS | 1 (bases 1 to 96599) | | | | |
| TITLE | Laird,G. | | | | |
| JOURNAL | Direct Submission | | | | |
| COMMENT | Submitted (31-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Jan 16, 2000 this sequence version replaced gi:6688397. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WormPDB; Information on the WormPDB database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpe/ the entire insert of clone RP1-309K20 This sequence was generated from part of bacterial clone contigs of human Chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chn20 | | | | |

| | | | |
|----|-------|---|-------|
| Db | 18832 | ACTGATATCCCAAGGAGGAGAG69TGGG69T6AGCCAGAGATTGTCCACTGCATCCAGCCT | 18773 |
| QY | 4891 | GAGTGAACAGGCTATCTC-----AAAAGCAACAAATAATAGTTTATCTAAACGGTA | 4641 |
| Db | 18772 | GAGCCACAGAGT6AGACTGTCTGTCTCAAAAAAAAAAAAAAAAAAGAAAACAAACACC | 18713 |
| QY | 4642 | AGGTATTAATACAGAAATATATGATAGCATTTTAAATTGAAAAAGCAATTAAATGATACAG | 4701 |
| Db | 18712 | ACATTTTGAAAGACAATAGCATGTATGTCAATGACACAAAGTTGCACAGCTAGGCCAAACAGTCC | 18653 |
| QY | 4702 | GATTGTAAAAATATCAATATACATAAATTTCTTGTCTTAATATATGCTAGCAACAAGCA | 4761 |
| Db | 18652 | AGGTAGAGGAATTAACACTGATCATCATTTGAATATCCCAATCTCTGCCCTGAGTGCAGTC | 18593 |
| QY | 4762 | CATTGGTTTTTACTAGGGCACCAGGTACTTTAAAAAAGTTAGGCGCAGCCACAGGGG | 4821 |
| Db | 18592 | CTCTTAATCTCCCATCTCACAACTCTTATGAATGTAGGAGAGCTGGGGCGGCGCATATGACGG | 18533 |
| QY | 4822 | CTCACACCTTATATCCAGCACTTTGGGAGGCCAGGAGGAGGATCACTTATAGCCACAG | 4881 |
| Db | 18532 | CTCACGCTGTACTCCAGACACTGTGTGAGACCCAGCA-GTGAGCACTTATAGGTCAAG | 18474 |
| QY | 4882 | AGTTTAGACCTGAGCAACATAGGAGATCTGATTTGTCTCTATAAAAAATTAATAAA | 4941 |
| Db | 18473 | AGTTCAAGACCAACTGCGCAATATATGTTGAACCTCATG--TCTACTAAATAATACAAAA | 18416 |
| QY | 4942 | TTGGCTAAGGCCCTTTGGGCTTACACCCGCTAATCCAGCACTTTGGGAGGCCGAGCGG | 4998 |
| Db | 18415 | TTTACCTGGGTGGGTGGCGGTGGTCCGTGTAGTTTCAACTTCTTTGGGAGGCTGTAGCGAG | 18359 |

RESULT 15
HSDJ689N3/C

| LOCUS | 121502 bp | DNA | linear | PRI 31-JAN-2002 |
|------------|---|-----|--------|-----------------|
| DEFINITION | Human DNA sequence from clone RP4-689N3 on chromosome Xp11.1-11.4 | | | |

Contains the 5' end of the RP2 gene for retinitis pigmentosa 2 (X-linked recessive), the 5' end of the SLC9A6 gene for solute carrier family 9 (sodium/hydrogen exchanger) isoform 6 and a CpG Island, complete sequence.

REVISION AL050307
 VERSION AL050307.13 GI:6562569
 KEYWORDS HTG; Cpg Island; retinitis pigmentosa; RP2; SLC9A6; solute carrier family 9.

| | |
|----------|--------------|
| SOURCE | human. |
| ORGANISM | Homo sapiens |

REFERENCE
1 (bases 1 to 121502)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Submitted (31-JAN-2002) Wellcome Trust Sanger Institute, Hinxton.

COMMENT

Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humnery@sanger.ac.uk. ISA clone requests: clonerequest@sanger.ac.uk on Dec 12, 1999 this sequence version replaced g1:6016948.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em1, EMBL; Sw1, SWISSPROT; Tr1, TREMBL; Wp1, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/Wormpep This sequence is the entire insert of clone RP4-689N3 This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group.

Further information can be found at <http://www.sanger.ac.uk/HGP/Chrx>

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (I.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. Rp1-669N3 is from the library RPCR-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/dacpac/home.htm>
VECTOR: pCYPAC2.

FEATURES

SOURCE

```

repeat_region 990. .1287
/note="MER82 repeat: matches 351. .652 of consensus"
repeat_region 1734. 1933
/note="MER82 repeat: matches 145. .354 of consensus"
repeat_region 3656. .3767
/note="LIMA4A repeat: matches 5428. .5539 of consensus"
repeat_region 4209. 4303
/note="LIMA4A repeat: matches 5539. .6165 of consensus"
repeat_region 11604. 11715
/note="56 copies 2 mer aa 60% conserved"
gene complement(11740. .34902)
/gene="Rp2"
complement(101n(<11740. .11854,17701. .18366,34640. .34902))

```

```

/gene="RP2"
/product="dj689N3.1 (retinitis pigmentosa 2 (X-linked recessive))"
/note="match: cDNAs: Em:AA007590
match: ESTs: Em:AL121304"
/evidence="not_experimental"
/complement={join(.11740..11854,17701..18366,34640..34741)}

```

CDS

```

/gene="RP2"
/note="match: proteins: Tr:075695"
/codon_start=1
/evidence=not_experimental
/product="dj689N3.1 (retinitis pigmentosa 2 (X-linked recessive))"
/protein_id="CAB82030.1"
/db_xref="GI:7320873"
/translation="MGCFESKRADKRDSPDENEEERPKOYSMPDREKVDKDFESG
LKDERTVGLPCTVAGQQLIIDDCENCNIYIIRDHSTVYIIDCTNCIILGCVKSGVF
RNCRCQCTLACQGFRRVDRCKRLLEFLCCAPPILESSNIRFCGFMYRYELAFQFEL
DAGSIFENNTSNIHDFPVRGSELVLLPDAVVDYPIPTTEELVAASVTEAN
SIVPSRQOROKSDSCSLVYLFAGDYITIANAKRLIDEMWKGFFLVQTRKVSNAKEE
AIVPRERAPDPIETPRNK"

```

```

Repeat_region      23885..24286
                    /note="201 copies 2 mer at 69% conserved"
Repeat_region      24288..24329
                    /note="21 copies 2 mer at 76% conserved"
Repeat_region      26621..27654
                    /note="11P2 repeat: matches 5316..6144"
Repeat_region      29352..29556
                    /note="L1M4 repeat: matches 6094..6294"
Repeat_region      29558..29626
                    /note="L1M4 repeat: matches 5984..6052"
                    complement(32542..32643)
Misc_feature        /gene="RP2"
                    /note="match: GSS: Em:AQ381314"
                    32641..33111
                    /note="match: GSS: Em:AQ453772"
                    34202..35135
                    /note="CPG island"
                    /evidence="not_experimental"
                    35440..35479
                    /note="20 copies 2 mer at 77% conserved"
                    38874..39187
                    /note="match: GSS: Em:AQ569300"

```


Db 73804 CCTGAAAGTAAAGAGTTCAAGACAGCTGGCCACATGGCAAAACCTGTCTACTATAA 73745
 Oy 3589 GCTGGGCTCCCTGACACAGCTCCTGGGACACTGGAGCTGCGGCTGGGGGCTGGGCA 3648
 Db 73744 AATACAAAAATTAAGCAGCGGCTGGTGTGTGCTGTAAATCCAGCTACTGGGAGGCT 73685
 Oy 3649 GAGTAAAGAGTCACTAGAGAGGCTGCTCAAGCCAGGCGACAGAGGTTTATGACACCT 3708
 Db 73684 GAG----- 73682
 Oy 3709 TCTTCACACCCAGAGAGACCCCTGGAGCCAGGCTGTGTGTGCTGGCCACTTACTGACC 3768
 Db 73681 -----GCAAGAGAAATCACTTGAACCTGGGAGGGGAGCGCGAGTGAGCGGAGAT 73632
 Oy 3769 TGTTTTACTGAATCCACACAGACTCATAGGCCCAATAGTACATTAATAAAGAGAGAGA 3828
 Db 73631 TGCACACCTGCACCTGACAGCTGGAGTAAAGAAATGAGCCCTGTCTCAAAAAA 73572
 Oy 3829 GAGAGAGAGAGAGAGAGATGAGTCTCACTGTGTGTGCTCCAGGCTGGTCTGAGACTCT 3888
 Db 73571 AAAGTACAGAGATAG-----GAGTCTGTGCTTTTCTCCAGGTTGCTTGAACCTCT 73518
 Oy 3889 AGGCTCAAGCAATCCCTGCTTACCTGCTCCAAAGGGGCTGGATTTACAGGTGTGAGCTA 3948
 Db 73517 GGCCTCAAGAGATTTCTTGGCTTGGCTTCCAAAGTCTGAGATTTACAGGTGTGAGCCA 73458
 Oy 3949 CTGCACCTGACCAACA-----CATGCTACTTTT 3978
 Db 73457 CCACATCTGGCCTGAAATATTTTTCACAGGCGCTGATTAATTAATTAATTA 73398
 Oy 3979 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 4038
 Db 73397 TCATTATTAATCATTAATTGACAGGAGTCTGTCTGTGTGCTGACAGGCTGGAGTGCAGT-G 73339
 Oy 4039 GGGCAATCTGGGCTCACTGTAACCTGTCTCCAGGTGCAAGCAATTTCTGCTGCTTAC 4098
 Db 73338 GGGGTATCTGAGCTATGCACTGCAACCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 73279
 Oy 4099 CCTCTGAGTATGCTGGAATTTATGACACACACACACAGGCGCTGCTAATTTTCTTTT 4158
 Db 73278 CTTCTGAGTATGCTGGGACATACAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 73222
 Oy 4159 TCTGTATTTTATGAGACAGGGTTTATCATGTTTGGCCAGGCTGTCTTGAACCCCTG 4218
 Db 73221 -----ATTTTATGAGAGATGGGTTTACACATGTTATCCAGGCTGTGCTGAACTCTG 73168
 Oy 4219 ACCTCAAGTATCCACCACCTGCGCTCCCAAAAGTGTGGATTACAGGTGTGACCCAC 4278
 Db 73167 AGCTCAGGCAATCCAGCGGCTGAGCTCCCAAAAGTGTGGATTACAGGTGTGACCCAC 73108
 Oy 4279 CATGCACAGCCACATGTAATTTTAAATTAATTTTAAATTAATTAATTTATCTAA 4338
 Db 73107 TGTGCTGGCCATGGGCTGATTAATTTTAAAGGCAAGTGA---AGAGTTGGGAAAAGA 73051
 Oy 4339 GGGCAGTATGAGTCTGCGCTGTATATCCACACTTTGAGGGGCCAAAGTGTGGGGA 4398
 Db 73050 GGGTGGGTGAGTGTGCTGATGCTGATATCCAGCTCTTTGGAGGCCAAGGAGGAAGA 72991
 Oy 4399 TCACTTAAAGCTGGGATTT-----CAGCGTGGCAACATAGTGAACCCCGTCTTAC 4452
 Db 72990 TCCCTTGAAGCCAAAGATTGAGACACAGCTGGGCAACATAGTGAACCCCGTCTTAC 72931
 Oy 4453 AAAAAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4512
 Db 72930 CAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 72871
 Oy 4513 AGCTGAGTGTGGGATGGCTGAAGCCTGTGAGGCTGAGGCTGAGTGAATGATGATCAC 4572
 Db 72870 GGGCAAGGTGAGAGAGACCTGAGTCCAGAGTTTGAAGCAGAGTGAAGTGTGATCAT 72811
 Oy 4573 ACCACTGCACTTCAAGCTGTGATGACAGGCTATCTCAAAAGCAAAATTAATGTTTATC 4632
 Db 72810 GCCACTGCACTTCAAGCTGTGATGACAGAGTGAAGTGTGCTCAAAAAA----- 72759

Oy 4633 TAAAGTAAAGTATATATCAGAAATATATGATATATTAATTAATTAATTAATTAATTAAT 4692
 Db 72758 -----AAAAAAGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 72707
 Oy 4693 GATTACATGATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4748
 Db 72706 TTACAGACATCTGATATCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 72647
 Oy 4749 TAGCAACAGGACATTTGTTTCTTACTAGGCGACCAAGTACTTTAAAAAAAGTTAGGG 4808
 Db 72646 TTGAAGGTGGGAAAGGAGAGAGTCTTTAATTAATTAATTAATTAATTAATTAATTAAT 72587
 Oy 4809 CCAGCCACAGGCGCTCACACTGTATATCCAGCACTTTGGAGGCCAAGGAGAGATC 4868
 Db 72586 GCCAGATGATGCTCATGCTGTATATCCAGCACTTTGGAGGCCAAGGAGATGAGAT 72527
 Oy 4869 ACTTGAGCCAGGAGTTAGGA---CCTGAGCAACATAGGAGAG-TGCTGATCTTGTCT 4923
 Db 72526 ATTGAGGTGAGGATTTGAGACACAGCTGCGCAACATGTTGAATCTCATCTTATTTA 72467
 Oy 4924 CTATTAATAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4983
 Db 72466 AATTAACAACAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 72407
 Oy 4984 GGGAGCCGAGGC 4996
 Db 72406 AGGAGCTGGGCC 72394

Search completed: June 8, 2003, 18:10:24
 Job time : 12664.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 09:29:22 ; Search time 448.2 Seconds

(without alignments)
11199.691 Million cell updates/sec

Title: US-09-622-964-2

Perfect score: 2229
Sequence: 1 caggaggtccaccagccta.....aaaaaaaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

N_Geneseq_101002:*

1: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 2229 | 100.0 | 2229 | 20 | AA221227 |
| 2 | 2013 | 90.3 | 2429 | 20 | AA221228 |
| 3 | 1247.2 | 56.0 | 1263 | 20 | AAV99722 |
| 4 | 871.2 | 39.1 | 1916 | 20 | AA221229 |
| 5 | 642.2 | 28.8 | 16125 | 20 | AA221226 |
| 6 | 642.2 | 28.8 | 16650 | 22 | ABAI4559 |
| 7 | 642.2 | 28.8 | 18530 | 22 | ABAI4557 |
| 8 | 640.6 | 28.7 | 18537 | 22 | ABAI4558 |
| 9 | 587.8 | 26.4 | 1717 | 24 | ABL89697 |

| | | | | | |
|----|-------|------|-------|----|-----------|
| 10 | 476.4 | 21.4 | 7108 | 22 | ABAI4556 |
| 11 | 314.8 | 14.1 | 2861 | 23 | ABLI0793 |
| 12 | 294.6 | 13.2 | 305 | 22 | ABAI1094 |
| 13 | 275.8 | 12.4 | 1526 | 23 | AA572796 |
| 14 | 252 | 11.3 | 1608 | 23 | ABLI2609 |
| 15 | 234.8 | 10.5 | 424 | 20 | AAV86976 |
| 16 | 218.6 | 9.7 | 10760 | 23 | ABLI0792 |
| 17 | 215.4 | 9.7 | 1345 | 22 | ABLI2637 |
| 18 | 215.4 | 9.7 | 1350 | 22 | AAV76848 |
| 19 | 199 | 8.9 | 2196 | 23 | ABLI16195 |
| 20 | 179.4 | 8.0 | 3592 | 23 | ABLI2636 |
| 21 | 175 | 7.9 | 1198 | 24 | ABK83853 |
| 22 | 175 | 7.9 | 1198 | 24 | ABN95676 |
| 23 | 175 | 7.9 | 1234 | 23 | AA567682 |
| 24 | 175 | 7.9 | 1279 | 23 | AA585655 |
| 25 | 175 | 7.9 | 2183 | 23 | AA591587 |
| 26 | 139.8 | 6.3 | 3785 | 23 | ABLI2608 |
| 27 | 116.8 | 5.2 | 2104 | 23 | AA572795 |
| 28 | 105.6 | 4.7 | 3592 | 23 | ABLI2636 |
| 29 | 100.8 | 4.5 | 392 | 22 | AAI20109 |
| 30 | 100.8 | 4.5 | 401 | 22 | AAI20109 |
| 31 | 99.8 | 4.5 | 527 | 22 | AAI17589 |
| 32 | 98 | 4.4 | 507 | 22 | AA539532 |
| 33 | 98 | 4.4 | 507 | 22 | AAK88431 |
| 34 | 70.2 | 3.1 | 1662 | 23 | AA570598 |
| 35 | 69.6 | 3.1 | 7505 | 23 | ABLI16194 |
| 36 | 66 | 3.0 | 7108 | 22 | ABAI4556 |
| 37 | 66 | 3.0 | 16125 | 20 | AA221226 |
| 38 | 66 | 3.0 | 16650 | 22 | ABAI4559 |
| 39 | 66 | 3.0 | 18530 | 22 | ABAI4557 |
| 40 | 66 | 3.0 | 18537 | 22 | ABAI4558 |
| 41 | 62.2 | 2.8 | 482 | 23 | AA572794 |
| 42 | 60 | 2.7 | 60 | 24 | ABN3610 |
| 43 | 54.4 | 2.4 | 65 | 24 | ABN57193 |
| 44 | 53.2 | 2.4 | 390 | 22 | ABAI4560 |
| 45 | 50.2 | 2.3 | 449 | 22 | ABAI85050 |

ALIGNMENTS

| | | |
|----------|--|---------------------------------------|
| RESULT 1 | AA221227 | standard; CDNA; 2229 BP. |
| ID | AA221227 | |
| XX | AA221227; | |
| AC | AA221227; | |
| XX | 22-NOV-1999 | (first entry) |
| DT | | |
| XX | Human CGICE short form CDNA sequence. | |
| DE | | |
| XX | CGICE: Best's macular dystrophy; mutation; diagnosis; detection; BMD; age-related macular dystrophy; ss. | |
| KW | | |
| XX | Homo sapiens. | |
| OS | | |
| XX | | |
| XX | Key | Location/Qualifiers |
| FT | CDS | 105..1862 |
| FT | | /*tag-8 |
| FT | | /product= "CGICE long form protein" |
| FT | | /transl_except= (pos:465..467,aa:Ser) |
| XX | MO9943695-A1. | |
| XX | 02-SEP-1999. | |
| XX | | |
| XX | 22-FEB-1999; | 99WO-US03790. |
| XX | | |
| XX | 25-FEB-1998; | 98US-0075941. |
| XX | 18-DEC-1998; | 98US-0112926. |
| XX | | |
| XX | (MERI) MERCK & CO INC. | |
| XX | (UYUP-) UNIV UPPSALA. | |

Human nervous syst
Drosophila melanog
Human nervous syst
DNA encoding novel
Drosophila melanog
EST clone BK112.
Drosophila melanog
Drosophila melanog
Human secreted pro
Drosophila melanog
Drosophila melanog
Human CDNA differe
Gene #2174 used to
DNA encoding novel
DNA encoding novel
DNA encoding novel
Drosophila melanog
Human breast cance
Human breast cance
Human breast cance
cDNA encoding nove
Human digestive sy
DNA encoding novel
Drosophila melanog
Human nervous syst
Human nervous syst
Human nervous syst
DNA encoding novel
DNA encoding novel
Mouse spliced tran
Human nervous syst
Human polynucleoti

QY 1801 CACTCAAGATCATGATGATCTTATTGGCCCTTGGAAAAAGGATGAAGACACTTCT 1860
DB 1801 CACTCAAGATCATGATGATCTTATTGGCCCTTGGAAAAAGGATGAAGACACTTCT 1860
QY 1861 AACCTCTCTCTTATGGGGATGCTTGGCCAGGAGCTCTCACCCTGTGTATACACGCA 1920
DB 1861 AACCTCTCTCTTATGGGGATGCTTGGCCAGGAGCTCTCACCCTGTGTATACACGCA 1920
QY 1921 GGACACTGATCCAGTACAGCCATACAGCTGTCCACAGTGAAGAGCTGTCTACACAG 1980
DB 1921 GGACACTGATCCAGTACAGCCATACAGCTGTCCACAGTGAAGAGCTGTCTACACAG 1980
QY 1981 CCTGAATCAAAATGGTATAGCTTAATAGATATAAAATCCAGACTCTTCACTTAATGCC 2040
DB 1981 CCTGAATCAAAATGGTATAGCTTAATAGATATAAAATCCAGACTCTTCACTTAATGCC 2040
QY 2041 TTTTATCTATTTAAATCTGCAAGCTAGACTGAACCTTGGAAACATTTAACTCAGACTC 2100
DB 2041 TTTTATCTATTTAAATCTGCAAGCTAGACTGAACCTTGGAAACATTTAACTCAGACTC 2100
QY 2101 TGGATTCAGAGTGGGAGACCTTATCTATCTGAATCAAGACAGCCACCTTAGTAT 2160
DB 2101 TGGATTCAGAGTGGGAGACCTTATCTATCTGAATCAAGACAGCCACCTTAGTAT 2160
QY 2161 ACTGCCCAACTATGATTTAATTAATACAAATCTCTTAAAAA 2220
DB 2161 ACTGCCCAACTATGATTTAATTAATACAAATCTCTTAAAAA 2220
QY 2221 AAAAAAAAA 2229
DB 2221 AAAAAAAAA 2229

RESULT 2
AA221228
ID AA221228 standard; cDNA; 2429 bp.

AC AA221228;

DT 22-NOV-1999 (first entry)

DE Human CG1CE long form cDNA sequence.

KM CG1CE; Best's macular dystrophy; mutation; diagnosis; detection;

KM BMD; age-related macular dystrophy; ss.

OS Homo sapiens.

FT Key Location/Qualifiers
FT CDS 105..1412
FT /product- "CG1CE short form protein"

PN MO9943695-A1.

PD 02-SEP-1999.

PF 22-FEB-1999; 99MO-US03790.

PR 25-FEB-1998; 98US-0075941.

PR 18-DEC-1998; 98US-0112926.

PA (MERI) MERCK & CO INC.

PI (UYUP-) UNIV UPPSALA.

PI Petrukhin K, Caskey CT, Metzker M, Wadelius C;

DR MPI, 1999-540560/45.

DR P-PSDB; AAY29954.

XX Human and mouse polynucleotides encoding CG1CE polypeptides

PS Claim 2; Fig 4; 67pp; English.

XX The present sequence represents the human CG1CE cDNA sequence, which
CC when mutated is responsible for Best's macular dystrophy (BMD).
CC Polynucleotides encoding CG1CE are useful for diagnosing whether a
CC patient carries a mutation in the CG1CE gene. Normal and mutated
CC CG1CE proteins are useful for identifying activators and/or inhibitors
CC of these proteins, in order to treat BMD. The CG1CE gene offers a
CC simpler and cheaper method of diagnosing BMD without the need for the
CC presence of the patient. The gene may also be useful to discovering
CC the genetic cause of age-related macular dystrophy.

SQ Sequence 2429 BP; 614 A; 694 C; 613 G; 508 T; 0 other;

Query Match 90.3%; Score 2013; DB 20; Length 2429;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 2226; Conservative 0; Mismatches 0; Indels 203; Gaps 1;

QY 1 CAGGAGTCCACACAGCTAGTGCAGACCTTCTGTGGATCATCGGACCCACTGGAA 60
DB 1 CAGGAGTCCACACAGCTAGTGCAGACCTTCTGTGGATCATCGGACCCACTGGAA 60
QY 61 CCCACCTGACCCAGGCGCAGCTGTGCAGGCGCCAGACCTTCTGTGGATCATCGGACCCACTGGAA 120
DB 61 CCCACCTGACCCAGGCGCAGCTGTGCAGGCGCCAGACCTTCTGTGGATCATCGGACCCACTGGAA 120
QY 121 CAAAGCAAGTGGCTAATGCCCGCTTAGGCTCTCTCCGCTCTGCTGTGCGGG 180
DB 121 CAAAGCAAGTGGCTAATGCCCGCTTAGGCTCTCTCCGCTCTGCTGTGCGGG 180
QY 181 GCAGCATCTCAAGCTCTATATGCGAGTCTTAATCTTCTGCTGTGCTACTACATCA 240
DB 181 GCAGCATCTCAAGCTCTATATGCGAGTCTTAATCTTCTGCTGTGCTACTACATCA 240
QY 241 TCCGCTTATTTATAGCTGCGCCCTCAGGAAAGCAAGACCTGATTTGAGAACTGA 300
DB 241 TCCGCTTATTTATAGCTGCGCCCTCAGGAAAGCAAGACCTGATTTGAGAACTGA 300
QY 301 CTCTGATTTGCGAGACGCTATACATCCAGCTATCCCATTTCTTCTGCTGTGCGCTTACG 360
DB 301 CTCTGATTTGCGAGACGCTATACATCCAGCTATCCCATTTCTTCTGCTGTGCGCTTACG 360
QY 361 TGACGCTGTGCTGACCCGCTGTGTGAAGCAAGTACGAAACCTTCTGCGCCACCGCC 420
DB 361 TGACGCTGTGCTGACCCGCTGTGTGAAGCAAGTACGAAACCTTCTGCGCCACCGCC 420
QY 421 TCATGAGCTGTGTGTGCGGCTGTGCGAAGCAAGCAAGGCGGCTGTGCGGC 480
DB 421 TCATGAGCTGTGTGTGCGGCTGTGCGAAGCAAGCAAGGCGGCTGTGCGGC 480
QY 481 GCAAGCTCATCCGCTACGCCAACCTGGGCAACGCTCATCTGCGAGCGTACGACCG 540
DB 481 GCAAGCTCATCCGCTACGCCAACCTGGGCAACGCTCATCTGCGAGCGTACGACCG 540
QY 541 CAGTCTACAAAGCGCTTCCCGAGCGCCAGACCTGTGCAAGCGGCTTATGACTCCGG 600
DB 541 CAGTCTACAAAGCGCTTCCCGAGCGCCAGACCTGTGCAAGCGGCTTATGACTCCGG 600
QY 601 CAGAACCAAGCAAGTGTGAGAAACTGAGCTACACCAACATGTTGTGGTGGCTGGG 660
DB 601 CAGAACCAAGCAAGTGTGAGAAACTGAGCTACACCAACATGTTGTGGTGGCTGGG 660
QY 661 TGTGTTTGGCAACCTGTCAATGAAGCGTGTGAGGCTGCAATCCGGGACCTATCC 720
DB 661 TGTGTTTGGCAACCTGTCAATGAAGCGTGTGAGGCTGCAATCCGGGACCTATCC 720
QY 721 TGTCTCAGAGCTGTCTGAAGAGATGAACACTTGTGCTACTCACTGTGAGACCTGTATG 780
DB 721 TGTCTCAGAGCTGTCTGAAGAGATGAACACTTGTGCTACTCACTGTGAGACCTGTATG 780
QY 781 CTTACGACCTGATAGTATGATCCCACTGTGTATACAGGCTGTGCTGTGAGGCTGTATCA 840
DB 781 CTTACGACCTGATAGTATGATCCCACTGTGTATACAGGCTGTGCTGTGAGGCTGTATCA 840

| | | | |
|----|------|---|------|
| QY | 84 | GCTTCCTCCGAGCTGTGCTAGTTGGGGGGGAGTTTCGAACCCGAGCCCAAGGCTCAACCTG | 900 |
| Db | 841 | GCTTCCTCCGAGCTGTGCTAGTTGGGGGGGAGTTTCGAACCCGAGCCCAAGGCTCAACCTG | 900 |
| QY | 901 | GCCATGAGCTGGAGACCTGTTGTGGCCGCTCTTCACGTTCTCGACGTTCTGTTCTATGTGG | 960 |
| Db | 901 | GCCATGAGCTGGAGACCTGTTGTGGCCGCTCTTCACGTTCTCGACGTTCTGTTCTATGTGG | 960 |
| QY | 961 | GCTGGCTGAA----- | 970 |
| Db | 961 | GCTGGCTGAAAGTGGGCTCTCTCCAGGGCCCTGCTGGGCTGGAGGCAATGGCCAGAGGGCTC | 1020 |
| QY | 971 | ----- | 970 |
| Db | 1021 | ATGGCCAGCAGCTCTTGAGACGAGGATGCAGTGTCAAGAAAGGAGGTCTCACGGGTAG | 1080 |
| QY | 971 | ----- | 970 |
| Db | 1081 | AAAGCAGCCAGGGCGTGGGGGCGACACACTGTAAATCCAGGCTACTCGGGAAGGCTGAGGCAG | 1140 |
| QY | 971 | -----GCTGGCAGAGCAGCTCATCAACCCCTT | 997 |
| Db | 1141 | GAGAAATCGCTTGAACCCGGGAGGGGAGGTTGTGTGGTGGCAGAGCAGCTCATCAACCCCTT | 1200 |
| QY | 998 | TGGAGAGGATGATGATGATATTTTGGACCACTGGATTTGTGCACAGGAAATTTGGAGGTGC | 1057 |
| Db | 1201 | TGGAGAGGATGATGATGATATTTTGGACCACTGGATTTGTGCACAGGAAATTTGGAGGTGC | 1260 |
| QY | 1058 | CCCTTTGGCTGTGGATGAGATGCACACAGGACCTGCTCGATGGAGCGGACATGTACTG | 1117 |
| Db | 1261 | CCCTTTGGCTGTGGATGAGATGCACACAGGACCTGCTCGATGGAGCGGACATGTACTG | 1320 |
| QY | 1118 | GAATTAAGCCGAGCCACAGCCCCCTACACAGCTGCTTCGGCCAGTTCCGTGAGGCTC | 1177 |
| Db | 1321 | GAATTAAGCCGAGCCACAGCCCCCTACACAGCTGCTTCGGCCAGTTCCGTGAGGCTC | 1380 |
| QY | 1178 | CTTATATGGGCTCCACGCTTCAACATCAGCGTGAACAAAGAGGAGATGGATCCAGCCCAA | 1237 |
| Db | 1381 | CTTATATGGGCTCCACGCTTCAACATCAGCGTGAACAAAGAGGAGATGGATCCAGCCCAA | 1440 |
| QY | 1238 | TCAGGAGGAGCAGGAGAGATGCTCAGCGTGGCATTTGGCCGCTTCCTAGGCGCTGCAGTC | 1297 |
| Db | 1441 | TCAGGAGGAGCAGGAGAGATGCTCAGCGTGGCATTTGGCCGCTTCCTAGGCGCTGCAGTC | 1500 |
| QY | 1298 | CCATGATTCACCATCTCCACGAGGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGGA | 1357 |
| Db | 1501 | CCATGATTCACCATCTCCACGAGGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGGA | 1560 |
| QY | 1358 | ATCCCTCTCTCCACGAGGGCTGCCCCAAAACCAACAAGGACGCAACAAGACGTTAGGGG | 1417 |
| Db | 1561 | ATCCCTCTCTCCACGAGGGCTGCCCCAAAACCAACAAGGACGCAACAAGACGTTAGGGG | 1620 |
| QY | 1418 | CCAGGAAGACAAAGGCGCTGGAAGCTTAAAGGCTGTGGAGCGCTTCAAGTGTGGCCACT | 1477 |
| Db | 1621 | CCAGGAAGACAAAGGCGCTGGAAGCTTAAAGGCTGTGGAGCGCTTCAAGTGTGGCCACT | 1680 |
| QY | 1478 | GTATCAGAGGCGCAGGCTACTACAGTGGCCCCACACAGGCGCCCTCGAGCCCACTCCCATGTT | 1537 |
| Db | 1681 | GTATCAGAGGCGCAGGCTACTACAGTGGCCCCACACAGGCGCCCTCGAGCCCACTCCCATGTT | 1740 |
| QY | 1538 | CTTCCCCCTAGAACCATATCAGCGCGGTCAAACTTTCACAGTGTCAAGGCAATGACACCCA | 1597 |
| Db | 1741 | CTTCCCCCTAGAACCATATCAGCGCGGTCAAACTTTCACAGTGTCAAGGCAATGACACCCA | 1800 |
| QY | 1598 | AGACAAAAGCTTAAAGCTGTGAAGTCTTGGGGGCCAAGAAAGTTTGAATGTGCTCAGA | 1657 |
| Db | 1801 | AGACAAAAGCTTAAAGCTGTGAAGTCTTGGGGGCCAAGAAAGTTTGAATGTGCTCAGA | 1860 |
| QY | 1658 | GAGCGATGGGCGCTTGATGAGACACCCAGAAGTATCTCAAGTGAAGAGAAACGTGGA | 1717 |
| Db | 1861 | GAGCGATGGGCGCTTGATGAGAGCAACCAAGAAGTATCTCAAGTGAAGAGAAACGTGGA | 1920 |
| QY | 1718 | GTTTAACTGAGGATATGCCAGAGATCCCGAAAAATCAGCTCAAGAACCTTTGGAACA | 1777 |

| Db | 1921 | GTTTTAACTGACGGATATGCCAGAGATCCCGAATAATACACTGCAAAAGACCTTTGGAAACA | 1980 |
|-----------------------------------|---|--|------|
| QY | 1778 | ATCACAACCAACATATACACACTACACTCAAGTCAAAAGATACATGGATCCTTATTGGGCTTGGGA | 1837 |
| Db | 1981 | ATCACAACCAACCAATACACACTATACACTCAAAAGATACATGGATCCTTATTGGGCTTGGGA | 2040 |
| QY | 1838 | AAAGAAGGATGAACACATTCCTTAACCTGGCTCTTAAATGGGAGTCTTGGCAGCCAGGT | 1897 |
| Db | 2041 | AAACAGGAGATGAACACATTCCTTAACCTGGCTCTTAAATGGGAGTCTTGGCAGCCAGGT | 2100 |
| QY | 1898 | CCTCAGCTGTGTGTACACACAGCAGACACTGATCTCAGTCAACAGCCATACAGCTGTCCACA | 1957 |
| Db | 2101 | CCTCAGCTGTGTGTACACACAGCAGACACTGATCTCAGTCAACAGCCATACAGCTGTCCACA | 2160 |
| QY | 1958 | CTGAAGAAGCGTGTCTTACACACAGCCTGAATCAAAATGTTAGCTTAAATAGATAAAAATCCC | 2017 |
| Db | 2161 | CTGAAGAAGCGTGTCTTACACACAGCCTGAATCAAAATGTTAGCTTAAATAGATAAAAATCCC | 2220 |
| QY | 2018 | AGACTACTCTGACCTTTAAAGCCTTTTATTCATATAAAACGTGAAGCTGACATGAACCA | 2077 |
| Db | 2221 | AGACTACTCTGACCTTTAAAGCCTTTTATTCATATAAAACGTGAAGCTGACATGAACCA | 2280 |
| QY | 2078 | TTGGAACATTTAACTAGACTGAGACTGAGATTCAGATCGGGAACCTTAGTCTATCTGAT | 2137 |
| Db | 2281 | TTGGAACATTTAACTAGACTGAGACTGAGATTCAGATTCGGAACCTTAGTCTATCTGAT | 2340 |
| QY | 2138 | CCAAGACAGCCACACCTTAGTATAGTCTGCCCAAACTAAATGAGTTTAAATATACAAATACT | 2197 |
| Db | 2341 | CCAAGACAGCCACACCTTAGTATAGTCTGCCCAAACTAAATGAGTTTAAATATACAAATACT | 2400 |
| QY | 2198 | CGTTAAAAAATTT | 2226 |
| Db | 2401 | CGTTAAAAAATTT | 2429 |
| RESULT 3 | | | |
| AAV99722 standard; cDNA: 1263 BP. | | | |
| XX | AAV99722; | | |
| AC | AAV99722; | | |
| XX | 26-APR-1999 | (first entry) | |
| DT | | | |
| XX | | | |
| DE | Human adult retina secreted protein bk112_15 | cDNA. | |
| XX | | | |
| KM | Secreted protein; human; retina; bk112_15; ds. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |
| FH | Key | Location/Qualifiers | |
| FT | CDS | 100..885 | |
| FT | | /*tag= a | |
| XX | | | |
| PN | W09856909-A2. | | |
| XX | | | |
| PD | 17-DEC-1998. | | |
| XX | | | |
| PF | 08-JUN-1998; | 98WO-US11822. | |
| XX | | | |
| PR | 05-JUN-1998; | 98US-0092722. | |
| XX | 11-JUN-1997; | 97US-0873218. | |
| XX | | | |
| PA | (GENY) GENETICS INST INC. | | |
| PI | Agostino MJ, Fechtel K, Howes SH, Jacobs K, Lavallie ER; | | |
| XX | McCoy JM, Racie LA, Spaulding V, Treacy M; | | |
| XX | WPI; 1999-080899/07. | | |
| DR | P-PSDB; AAM95345. | | |
| PT | New polynucleotides encoding secreted human proteins - derived from human foetal brain, adult testes, foetal kidney, adult thyroid or | | |

PR 18-DEC-1998; 98US-0112926.
 XX (MERI) MERCK & CO INC.
 PA (UYUP-) UNIV UPPSALA.
 XX Petrukhin K, Caskey CT, Metzker M, Madellius C;
 XX WPI: 1999-540560/45.
 DR P-PSDB; AAY29955.
 XX Human and mouse polynucleotides encoding CGICE polypeptides
 PS Claim 2; Fig 8; 67pp; English.
 XX The present sequence represents the mouse CGICE cDNA sequence, which
 CC when mutated is responsible for Best's macular dystrophy (BMD).
 CC Polyucleotides encoding CGICE are useful for diagnosing whether a
 CC patient carries a mutation in the CGICE gene. Normal and mutated
 CC CGICE proteins are useful for identifying activators and/or inhibitors
 CC of these proteins, in order to treat BMD. The CGICE gene offers a
 CC simpler and cheaper method of diagnosing BMD without the need for the
 CC presence of the patient. The gene may also be useful to discovering
 CC the genetic cause of age-related macular dystrophy.
 XX
 SQ Sequence 1916 BP; 514 A; 499 C; 457 G; 446 T; 0 other:
 Query Match 39.1%; Score 871.2; DB 20; Length 1916;
 Best Local Similarity 77.1%; Pred. No. 1.7e-193;
 Matches 1116; Conservative 0; Mismatches 293; Indels 39; Gaps 3;

QY 756 CGTACTAGTGTGGACACCTGTATGCTACGACTGATTAATCCCACTGGTATACA 815
 DB 662 CGTACTAGTGTGGACACCTGTATGCTACGACTGATTAATCCCACTGGTATACA 721
 QY 816 CAGGTGGTACGTGGCGGTGATACAGCTTCTCTGACTGTCTGATGGGCGGAGTT 875
 DB 722 CAGGTGGTACGTGGCGGTGATACAGCTTCTCTGACTGTCTGATGGGCGGAGTT 781
 QY 876 CTGAACCCAGCCAGGCTTACCTGGCCATGAGCTGAGCTGATGCTGCTTACG 935
 DB 782 CTGAACCCAGCCAGGCTTACCTGGCCATGAGCTGAGCTGATGCTGCTTACG 841
 QY 936 TTCTGCACTTCTTCTATGTTGGCTGGCTGGAAGTGGCAGACAGCTATCAACCC 995
 DB 842 ATCCGCAATTCCTTATGCTTGGCTGGCTGGAAGTGGCAGACAGCTATCAACCC 901
 QY 996 TTGGAGAGGATGATGATGATTTTGAACAAGCTGATGCTGAGAGAAATTTGAGTG 1055
 DB 902 TTGGAGAGGATGATGATGATTTTGAACAAGCTGATGCTGAGAGAAATTTGAGTG 961
 QY 1056 TTCCGTTGGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1115
 DB 962 TTCCGTTGGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021
 QY 1116 TGGATTAAGCCGAGACACAGCCCTTACACAGCTGCTCCGCAATTCGTCGAGGC 1175
 DB 1022 TGGATTAAGCCGAGACACAGCCCTTACACAGCTGCTCCGCAATTCGTCGAGGC 1081
 QY 1176 TCCCTTAAGCCGAGACACAGCCCTTACACAGCTGCTCCGCAATTCGTCGAGGC 1235
 DB 1082 TCCCTTAAGCCGAGACACAGCCCTTACACAGCTGCTCCGCAATTCGTCGAGGC 1141
 QY 1236 AATCAG 1283
 DB 1142 AATCAG 1201
 QY 1284 CTAGGCTGAG 1343
 DB 1202 CTAGGCTGAG 1261
 QY 1344 TGGCCCAAG 1403
 DB 1262 TGGCCCAAG 1315
 QY 1404 CAGAACGTTAGGAG 1463
 DB 1316 CAGAACGTTAGGAG 1354
 QY 1464 AAGTGGGCGGCTATATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1523
 DB 1354 AAGTGGGCGGCTATATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1414
 QY 1524 CCCACTCC 1531
 DB 1415 CACCTTAC 1422

RESULT 5
 AA221226
 ID AA221226 standard; DNA; 16125 BP.
 XX AA221226;
 AC 22-NOV-1999 (first entry)
 XX
 DE Human CGICE genomic DNA sequence.
 XX CGICE: Best's macular dystrophy; mutation; diagnosis; detection;
 KW BMD; age-related macular dystrophy; ss.
 XX Homo sapiens.
 OS

| PH | Key | Location/Qualifiers |
|----|--------|---------------------------|
| PT | Intron | 1..1513 |
| PT | | /*tag= a |
| PT | | /number= 1 |
| PT | exon | 1514..1570 |
| PT | | /*tag= b |
| PT | | /number= 1 |
| PT | Intron | 1571..3278 |
| PT | | /*tag= c |
| PT | | /number= 2 |
| PT | exon | 3279..3466 |
| PT | | /*tag= d |
| PT | | /number= 2 |
| PT | CDS | 3335..15787 |
| PT | | /*tag= e |
| PT | | /product= "CGICE protein" |
| PT | | /note= "contains introns" |
| PT | Intron | 3467..6616 |
| PT | | /*tag= f |
| PT | | /number= 3 |
| PT | exon | 6617..6711 |
| PT | | /*tag= g |
| PT | | /number= 3 |
| PT | Intron | 6712..7227 |
| PT | | /*tag= h |
| PT | | /number= 4 |
| PT | exon | 7228..7461 |
| PT | | /*tag= i |
| PT | | /number= 4 |
| PT | Intron | 7462..8353 |
| PT | | /*tag= j |
| PT | | /number= 5 |
| PT | exon | 8354..8508 |
| PT | | /*tag= k |
| PT | | /number= 5 |
| PT | | 8509..8896 |
| PT | | /*tag= l |
| PT | | /number= 6 |
| PT | exon | 8897..8974 |
| PT | | /*tag= m |
| PT | | /number= 6 |
| PT | Intron | 8975..9676 |
| PT | | /*tag= n |
| PT | | /number= 7 |
| PT | exon | 9677..10032 |
| PT | | /*tag= o |
| PT | | /number= 7 |
| PT | Intron | 10033..11212 |
| PT | | /*tag= p |
| PT | | /number= 8 |
| PT | exon | 11213..11293 |
| PT | | /*tag= q |
| PT | | /number= 8 |
| PT | Intron | 11294..11556 |
| PT | | /*tag= r |
| PT | | /number= 9 |
| PT | exon | 11557..11708 |
| PT | | /*tag= s |
| PT | | /number= 9 |
| PT | Intron | 11709..13918 |
| PT | | /*tag= t |
| PT | | /number= 10 |
| PT | exon | 13919..14557 |
| PT | | /*tag= u |
| PT | | /number= 10 |
| PT | Intron | 14558..15768 |
| PT | | /*tag= v |
| PT | | /number= 11 |
| PT | exon | 15769..16125 |
| PT | | /*tag= w |
| PT | | /number= 11 |
| XX | | |

| | |
|---------------------------|--|
| XX | 02-SEP-1999. |
| PD | |
| XX | |
| PF | 22-FEB-1999; 99WO-US03790. |
| XX | |
| PR | 25-FEB-1998; 98US-0075941. |
| PR | 18-DEC-1998; 98US-0112926. |
| XX | |
| PA | (MERI) MERCK & CO INC. |
| PA | (UYUP-) UNIV UPESALA. |
| PI | Petrukhin K, Caskey CT, Metzker M, Wadellius C; |
| XX | |
| XX | WPI, 1999-540560/45. |
| DR | P-PsDB; AAY29953. |
| PT | |
| PS | Human and mouse polynucleotides encoding CGICE polypeptides - |
| XX | |
| PS | Claim 2; Fig 1; 67pp; English. |
| XX | |
| CC | The present sequence represents the human CGICE gene, which when |
| CC | mutated is responsible for Best's macular dystrophy (BMD). |
| CC | Polynucleotides encoding CGICE are useful for diagnosing whether a |
| CC | patient carries a mutation in the CGICE gene. Normal and mutated |
| CC | CGICE proteins are useful for identifying activators and/or inhibitors |
| CC | of these proteins, in order to treat BMD. The CGICE gene offers a |
| CC | simpler and cheaper method of diagnosing BMD without the need for the |
| CC | presence of the patient. The gene may also be useful to discovering |
| CC | the genetic cause of age-related macular dystrophy. |
| XX | |
| SQ | Sequence 16125 BP; 3988 A; 4175 C; 4215 G; 3726 T; 21 other; |
| Query Match | 28.8%; Score 642.2; DB 20; Length 16125; |
| Best Local Similarity | 99.5%; Pred. No. 1.3e-139; |
| Matches 644; Conservative | 0; Mismatches 3; Indels 0; Gaps 0; |
| OY | 1202 CAGCTGTAACAAAGAGAGATGGAGTTTCAGGCCCAATCAGAGAGCAGAGAGATGCTCA 1261 |
| Db | 13916 CAGCTGTAACAAAGAGAGATGGAGTTTCAGGCCCAATCAGAGAGCAGAGAGATGCTCA 13979 |
| OY | 1262 CGGTGCATCATTTGGCGCGTTCCTTAGGCGCTGCAGTCCCATATACCATTCCTCCAGGGC 1321 |
| Db | 13976 CGGTGCATCATTTGGCGCGTTCCTTAGGCGCTGCAGTCCCATATACCATTCCTCCAGGGC 14039 |
| OY | 1322 AAACCTAAGAGACCCTACTCTGTGGCCCAAGAGGAATCCCTTGCCAGAGGGCGCTGCC 1381 |
| Db | 14036 AAACCTAAGAGACCCTACTCTGTGGCCCAAGAGGAATCCCTTGCCAGAGGGCGCTGCC 14099 |
| OY | 1382 CAAAAACAAGAGAGCCCAACAGAACGTTAGGGGCGAGAGAGACAACAAGAGGCTGGNA 1441 |
| Db | 14096 CAAAAACAAGAGAGCCCAACAGAACGTTAGGGGCGAGAGAGACAACAAGAGGCTGGNA 14155 |
| OY | 1442 GCTTAAGGCTGTGGAGAGGCTTCAAGTGTGGCCCACTGTATCAGAGGCCAGGCTACTACAG 1501 |
| Db | 14156 GCTTAAGGCTGTGGAGAGGCTTCAAGTGTGGCCCACTGTATCAGAGGCCAGGCTACTACAG 14215 |
| OY | 1502 TTCGCCACAGAGAGGCCCTCAGAGCCCACTGCCATGTTCTTCCCCTAGAACAATCAGAGGCC 1561 |
| Db | 14216 TTCGCCACAGAGAGGCCCTCAGAGCCCACTGCCATGTTCTTCCCCTAGAACAATCAGAGGCC 14275 |
| OY | 1562 GTCAAAGCTTCAAGTGTACAGAGGCATATAGACCAACAAGAGAAAAGCTTAAAGACTGTAG 1621 |
| Db | 14276 GTCAAAGCTTCAAGTGTACAGAGGCATATAGACCAACAAGAGAAAAGCTTAAAGACTGTAG 14335 |
| OY | 1622 TTCTGGGGGCCAAGAAAAGTTTTGAATTGCTCTCAGAGAGCGATGGGCCCTTATGAGAGCA 1681 |
| Db | 14336 TTCTGGGGGCCAAGAAAAGTTTTGAATTGCTCTCAGAGAGCGATGGGGCCTTATGAGAGCA 14395 |
| OY | 1682 CCCAGAAATATCTCAAGTGTAGAGAGAAAAGCTGTGAGTTTAACCTGACGGATATGCCAGA 1741 |
| Db | 14396 CCCAGAAATATCTCAAGTGTAGAGAGAAAAGCTGTGAGTTTAACCTGACGGATATGCCAGA 14455 |
| OY | 1742 GATCCCGAATAATCACTTCAAGAAACCTTTGGAACATCACCAACCAACATACACTAC 1801 |

Db 14456 GATCCCGGAATATCACTCAAGACCTTTGGAAACAATCACCACCAATACACACTAC 14515
QY 1802 ACTCAAGATCATCATGATCTTATTTGGCCCTTGGAAAAAGGATG 1848
Db 14516 ACTCAAGATCATCATGATCTTATTTGGCCCTTGGAAAAAGGATCTG 14562
RESULT 6
ABAL4559
ID ABAL4559 standard; DNA; 16650 BP.
XX
AC ABAL4559;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6890.
XX
KM Human; neotropic; neuroprotective; cytostatic; dermatological; virucide;
KM immunosuppressive; anti-infective; anti-HIV; antibacterial; vulnary;
KM antiparkinsonian; antischistosomal; antitubercular; cancer;
KM antitrematode; hepatotropic; cerebroprotective; anti-inflammatory;
KM antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KM antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KM neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WC200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.

PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.

PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249268.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI: 2001-541565/60.

Nucleic acids encoding 3224 human nervous system antigen polypeptides,
useful for preventing, diagnosing and/or treating nervous system
cancers and metastases -
XX
PS Disclosure; SEQ ID NO 6888; 1701pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABAI1004-ABR21534) and proteins
CC (ABBI678-ABBI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's

CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC XX

SO Sequence 18530 BP; 4512 A; 4799 C; 4984 G; 4235 T; 0 other;

Query Match 28.8%; Score 642.2; DB 22; Length 18530;
 Best Local Similarity 99.5%; Pred. No. 1.4e-139;
 Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1202 CAGCTGAACAAGAGAGATGAGTTCACGCCAATCAGAGAGAGAGATGCTCA 1261
 DB 11895 CAGCTGAACAAGAGAGATGAGTTCACGCCAATCAGAGAGAGATGCTCA 11954
 QY 1262 CGCTGGCATCATGCGCTTCTGAGGCTGAGTCCCATGATCATCTCTCCAGGGC 1321
 DB 11955 CGCTGGCATCATGCGCTTCTGAGGCTGAGTCCCATGATCATCTCTCCAGGGC 12014
 QY 1322 AAATCAAGAGACCAACTACTGTGGCCCAAGAGAGATCCCTTCCAGAGGGCTGCC 1381
 DB 12015 AAATCAAGAGACCAACTACTGTGGCCCAAGAGAGATCCCTTCCAGAGGGCTGCC 12074
 QY 1382 CAAAAACCAAG 1441
 DB 12075 CAAAAACCAAG 12134
 QY 1442 GCTTAAGGCTGTGAGAGGCTTCAAGTCTGAGGCTGAGTCTGAGAGAGAGAGAG 1501
 DB 12135 GCTTAAGGCTGTGAGAGGCTTCAAGTCTGAGGCTGAGTCTGAGAGAGAGAGAG 12194
 QY 1502 TGGCCCAAG 1561
 DB 12195 TGGCCCAAG 12254
 QY 1562 GTCAAAAGCTTCACAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
 DB 12255 GTCAAAAGCTTCACAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 12314
 QY 1622 TTCTGGGGCCAGAGAAAGTTTGAATGCTCTGAGAGAGAGAGAGAGAGAGAGAG 1681
 DB 12315 TTCTGGGGCCAGAGAAAGTTTGAATGCTCTGAGAGAGAGAGAGAGAGAGAGAG 12374
 QY 1682 CCCAGAAAGTATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1741
 DB 12375 CCCAGAAAGTATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 12434
 QY 1742 GATCCCCCGAAGAAATCAGTCAAGAACTTTGGAACAATCAGCAACAGATACACTAC 1801
 DB 12435 GATCCCCCGAAGAAATCAGTCAAGAACTTTGGAACAATCAGCAACAGATACACTAC 12494
 QY 1802 ACTCAAGATCAGATGATGATCTTATTGGGCTTGGAAACAGAGAGAG 1848
 DB 12495 ACTCAAGATCAGATGATGATCTTATTGGGCTTGGAAACAGAGAGAG 12541

RESULT 8
 ABA14558
 ID ABA14558 standard; DNA; 18537 BP.
 AC ABA14558;
 XX

XX 23-JAN-2002 (first entry)
 DT
 DE

DE Human nervous system related polynucleotide SEQ ID NO 6889.
 XX

KW Human; nootropic; neuroprotective; cytosolic; dermatological; virucide;
 KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnery;
 KW antiparkinsonian; antistoking; antinaemic; antiarthritic; cancer;

KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antileuc; anticonvulsant; antifungal;
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 XX Homo sapiens.
 OS
 PN WO200159063-A2.
 PD 16-AUG-2001.
 XX
 PF 17-JAN-2001; 2001MO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0217496.
 PR 26-JUL-2000; 2000US-0220863.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226881.
 PR 22-AUG-2000; 2000US-0226886.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231245.
 PR 08-SEP-2000; 2000US-0231411.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.

XX 17-JAN-2001; 2001WO-US01334.
 PF
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-MAR-2000; 2000US-0198123.
 PR 19-MAR-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217467.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231245.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 XX

PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249224.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-02359678.
 XX
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-541565/60.
 DX

PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 useful for preventing, diagnosing and/or treating nervous system
 cancers and metastases -
 PT
 XX
 PS Disclosure: SEQ ID NO 6887; 1701pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (AB11004-AB21533) and proteins
 CC (AB14678-AB218001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX
 SQ Sequence 7108 BP; 1632 A; 1882 C; 1989 G; 1605 T; 0 other:
 Query Match 21.4%; Score 476.4; DB 22; Length 7108;
 Best Local Similarity 99.4%; Pred. No. 5.1e-101;
 Matches 499; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
 QY 1202 CAGCCTGACAAAGAGAGATGAGTTCACCCCAATGAGAGAGAGAGATGCTCA 1261
 DB 6609 CAGCCTGACAAAGAGAGATGAGTTCACCCCAATGAGAGAGAGAGATGCTCA 1261
 QY 1262 CGCTGAGCATATTGGCCCTCTAGAGCTGACATGATCCATATGATCCATCCAGGGC 1321
 DB 6668 CGCTGAGCATATTGGCCCTCTAGAGCTGACATGATCCATATGATCCATCCAGGGC 1321
 QY 1322 AACTCAAGAGACCAACTACTGTGTCGCCCAAGAGAGAGAGAGAGAGAGAGAGAG 1381
 DB 6728 AACTCAAGAGACCAACTACTGTGTCGCCCAAGAGAGAGAGAGAGAGAGAGAGAG 1381
 QY 1382 CAAAGACCAAG 1441
 DB 6787 CAAAGACCAAG 1441
 QY 1442 GCTTAAAGCTGTGAGAGCTTCAAGTCTGAGCCCACTGATCAGAGAGAGAGAGAG 1501
 DB 6847 GCTTAAAGCTGTGAGAGCTTCAAGTCTGAGCCCACTGATCAGAGAGAGAGAGAG 1501
 QY 1502 TGGCCCAAG 1561
 DB 6907 TGGCCCAAG 1561
 QY 1562 GTCAAGAGCTCAAGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
 DB 6967 GTCAAGAGCTCAAGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
 QY 1622 TTTCTGGGCGCAAGAAAGTTTGAATTTGCTCTCAGAGAGAGAGAGAGAGAGAGAG 1681
 DB 7027 TTTCTGGGCGCAAGAAAGTTTGAATTTGCTCTCAGAGAGAGAGAGAGAGAGAGAG 1681
 QY 1682 CCCAGAGATATCTCAAGTGAAG 1703
 DB 7087 CCCAGAGATATCTCAAGTGAAG 7108
 RESULT 11
 ABL10793
 ID ABL10793 standard; cDNA; 2861 BP.
 XX
 AC ABL10793;

XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 26861.
 KW
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 OS
 OS Drosophila melanogaster.
 PN
 PN WO200171042-A2.
 PD
 PD 27-SEP-2001.
 PF
 PF 23-MAR-2001; 2001WO-US09231.
 PR
 PR 23-MAR-2001; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PERK) PE CORP NY.
 XX
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 DR WPI; 2001-656860/75.
 DR P-PSDB; ABB66690.
 XX
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 PS Claim 1; SEQ ID NO 26861; 21pp + Sequence Listing; English.
 PS
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB72072).
 CC
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX
 SQ Sequence 2861 BP; 761 A; 701 C; 723 G; 676 T; 0 other:
 Query Match 14.1%; Score 314.8; DB 23; Length 2861;
 Best Local Similarity 57.4%; Pred. No. 2.1e-63;
 Matches 656; Conservative 0; Mismatches 462; Indels 24; Gaps 4;
 QY 103 CCATGACATACCTTACACAGCAAGTGG---CTAATGCCCGCTTAGGCTCTCTCC 159
 DB 413 CAAATGACAAATTAATCTACACAGTGAAGTGGCGCTTGGCGCTTGGCTTTCTCA 472
 QY 160 GCGTGTCTGTGCTGGCGGGGAGACATCTCAAGCTGCTATAGCGAGTCTTAATCT 219
 DB 473 AATGCTGTCTGATGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 532
 QY 220 TCTGTCTGTCTACTACATATCCGCTTATTTAATGAGCTGGCCCTCAAGAGAACAC 279
 DB 533 TCTGTGACCAATTTACTATGACATCAACATGATGATGCTTGGCTCAACCCGACANA 592
 QY 280 AGCTGATGTTTGAAGAACTGACCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 339
 DB 593 AAGAAACCTTTGAGGCGCATTTGATGATGATGATGATGATGATGATGATGATGATG 652
 QY 340 CCTGTGTGTGGGCTTCTAGTGAAGCTGTGCTGAGCCGCTGAGGAGGAGGAGGAGGAG 399
 DB 653 CCTGTGTGTGTGTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 712
 QY 400 ACCTGCGTGGGCGGAGCGGCTCATGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 459
 DB 713 CCATTCCTGGCCAGATCCATCCATGCGCGTGTGTGTGACCTGCAATGTCATAGGCGAGATG 772

XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI: 2001-639362/73.
 XX DR P-PSDB: ABB08609.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 PS Claim 1; SEQ ID NO 8600; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcr_sequences.
 XX
 XX SQ Sequence 1526 BP; 404 A; 408 C; 357 G; 357 T; 0 other;
 Query Match 12.4%; Score 275.8; DB 23; Length 1526;
 Best Local Similarity 65.2%; Pred. No. 2.1e-54;
 Matches 406; Conservative 0; Mismatches 217; Indels 0; Gaps 0;
 QY 586 GCTTATGACTCCGCGCAGACACAGACAGTGGAGAACTGAGCTTACACACACATGT 645
 DB 1 GTTTATGACACAGATGAGAAAGAAATTTATTCACCACTCAATCTCCATCTGTAAT 60
 QY 646 TCTGGTGGCCCTGGGTGGTGGTCCCACTGTCATGAAAGGCGCTGGAGGTGAA 705
 DB 61 ATTTGGGTCATTCATCTGTTGGAAATCTTGGCAATCCGGAATGAAAGTAGAA 120
 QY 706 TCCGGAGCCTTATCTGCTCCAGAGCTCGTGAAGAGATGAAACCTGGCTACTAGT 765
 DB 121 TAGAGACATGTTGATCTGATCTGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 766 GTGAGACCTGTATGCTTACAGATGATGATGATGATGATGATGATGATGATGATGAT 825
 DB 181 GCAGCCTCTATTCGTTATGACGTGGGTGGATTCGCGGTGTATACACCAAGTTGCA 240
 QY 826 CTGTGGGGGTGTACAGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 885
 DB 241 CTCTGCTGCTGTATACCTCTCTCTTTCGCGGTGATGATGATGATGATGATGATGATGAT 300

QY 886 CCAAGCCTACCTGGCCATGAGCTGACCTGCTTGTGCGGCTTCAAGTTCAGCT 945
 DB 301 CCAAGGCTACGACAGGCGATCTTGGATCTTCAATTCCTTCAACCCCTTCAAT 360
 QY 946 TCTTCTTATGTTGGCTGGCTGAGTGGAGAGACAGCTCAATCAACCCCTTGGAGAG 1005
 DB 361 TCTTCTTATGAGGATGCTTAAAGTAGAGAGACAGCTTATCAACCCCTTGGAGAG 420
 QY 1006 ATGATGATGATTTGAGACCAACTGAGATTGTCGACAGAAATTTGAGGTGCTGTTGG 1065
 DB 421 ATGATGATGATTTGAGAACTAAGTGTGATGAGAAATTTGAGGTGCTGTTTAA 480
 QY 1066 CTGTGATGATGATGACAGACAGCTGCTGAGATGAGACCGGACATGATGAGAAATGAGC 1125
 DB 481 CTGTGAGCAATATGACATGAGCTTACCAAGATGAGAAAGACATTTACTGAGGAGATT 540
 QY 1126 CCGAGCCACAGCCCCCTACACAGCTGCTCCGCCACTTCCGTGAGCCTTCCTTATG 1185
 DB 541 CTGCTGCTCGGCCACATACATGATGAGCTGCTGATGATGATGATGATGATGATGATGAT 600
 QY 1186 GCTCCACTTCAACATCAGCTG 1208
 DB 601 GGTCAACAGTCCAGATGGGCTG 623
 RESULT 14
 ABL12609
 ID ABL12609 standard; cDNA; 1608 BP.
 XX
 XX ABL12609;
 AC
 AC 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 32309.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 PD 27-SEP-2001.
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 XX
 XX WPI: 2001-656860/75.
 DR P-PSDB: ABB68506.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS
 PS Claim 1; SEQ ID NO 32309; 21pp + Sequence Listing; English.
 CC
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes in Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO

Db 141 GATATGCCAGAGATCCCCGAAAAATCACCCTTGAAGAACCTTTGGAAACATCACCACCAAC 200
QY 1791 ATACACCTACACTCAAAAGATCACATGATCTTATTTGGGCTTGGAAACAGGATG 1848
Db 201 ATACACACTACACTCAAAAGATCACATGATCTTATTTGGGCTTGGAAACAGGCTG 258

Search completed: June 8, 2003, 11:23:47
Job time : 452.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 10:29:42 ; Search time 91.8557 Seconds
(Without alignments)
7441.920 Million cell updates/sec

Title: US-09-622-964-2
Perfect score: 2229
Sequence: 1 cagggaggtccaccagccta.....aaaaaaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY_MNC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/lna/5A.COMB.seq: *
2: /cgn2_6/prodata/1/lna/5B.COMB.seq: *
3: /cgn2_6/prodata/1/lna/6A.COMB.seq: *
4: /cgn2_6/prodata/1/lna/6B.COMB.seq: *
5: /cgn2_6/prodata/1/lna/PCMTUS.COMB.seq: *
6: /cgn2_6/prodata/1/lna/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 41.6 | 1.9 | 1212 | 4 | US-09-182-145-34 |
| 2 | 41.6 | 1.9 | 1212 | 4 | US-09-182-145-35 |
| 3 | 41.2 | 1.8 | 1161 | 1 | US-08-086-439C-2 |
| 4 | 41.2 | 1.8 | 1161 | 1 | US-08-434-877-2 |
| 5 | 41.2 | 1.8 | 1367 | 3 | US-08-475-742-3 |
| 6 | 41.2 | 1.8 | 1370 | 1 | US-08-056-051-1 |
| 7 | 41.2 | 1.8 | 1370 | 1 | US-07-928-611-17 |
| 8 | 41.2 | 1.8 | 1370 | 2 | US-08-487-811A-17 |
| 9 | 41.2 | 1.8 | 1370 | 4 | US-09-060-694-17 |
| 10 | 41.2 | 1.8 | 1370 | 4 | US-09-378-074-17 |
| 11 | 41.2 | 1.8 | 1370 | 5 | PCR-US93-07370-17 |
| 12 | 41.2 | 1.8 | 1466 | 1 | US-08-056-051-3 |
| 13 | 41.2 | 1.8 | 1466 | 1 | US-07-928-611-19 |
| 14 | 41.2 | 1.8 | 1466 | 2 | US-08-487-811A-19 |
| 15 | 41.2 | 1.8 | 1466 | 4 | US-09-060-694-19 |
| 16 | 41.2 | 1.8 | 1466 | 4 | US-09-378-074-19 |
| 17 | 41.2 | 1.8 | 1466 | 5 | PCR-US93-07370-19 |
| 18 | 41.2 | 1.8 | 1507 | 4 | US-09-453-323-1 |
| 19 | 41.2 | 1.8 | 1610 | 1 | US-08-056-051-5 |
| 20 | 41.2 | 1.8 | 1610 | 1 | US-07-928-611-21 |
| 21 | 41.2 | 1.8 | 1610 | 2 | US-08-487-811A-21 |
| 22 | 41.2 | 1.8 | 1610 | 4 | US-09-060-694-21 |
| 23 | 41.2 | 1.8 | 1610 | 4 | US-09-378-074-21 |
| 24 | 41.2 | 1.8 | 1610 | 5 | PCR-US93-07370-21 |
| 25 | 41.2 | 1.8 | 1700 | 3 | US-08-897-340-4 |
| 26 | 41.2 | 1.8 | 1700 | 3 | US-09-352-329-4 |
| 27 | 41.2 | 1.8 | 7218 | 1 | US-08-232-463-14 |

| | | | | | | |
|----|------|-----|------|---|-------------------|-------------------|
| 28 | 40.8 | 1.8 | 1194 | 4 | US-08-765-907A-9 | Sequence 9, App1 |
| 29 | 40.8 | 1.8 | 1233 | 4 | US-08-765-907A-12 | Sequence 12, App1 |
| 30 | 40.8 | 1.8 | 1561 | 4 | US-08-765-907A-11 | Sequence 11, App1 |
| 31 | 40.8 | 1.8 | 3124 | 4 | US-09-734-030-1 | Sequence 1, App1 |
| 32 | 40.8 | 1.8 | 4496 | 4 | US-08-765-907A-6 | Sequence 6, App1 |
| 33 | 40.6 | 1.8 | 2634 | 4 | US-09-463-238-3 | Sequence 3, App1 |
| 34 | 40.6 | 1.8 | 3275 | 4 | US-09-370-838-151 | Sequence 151, App |
| 35 | 40.4 | 1.8 | 1512 | 2 | US-08-909-965C-8 | Sequence 5, App1 |
| 36 | 40.4 | 1.8 | 1697 | 4 | US-09-346-408-5 | Sequence 5, App1 |
| 37 | 40.2 | 1.8 | 593 | 1 | US-08-524-757-1 | Sequence 262, App |
| 38 | 40.2 | 1.8 | 976 | 2 | US-08-504-459-9 | Sequence 9, App1 |
| 39 | 40.2 | 1.8 | 1373 | 5 | PCR-US92-02977-6 | Sequence 6, App1 |
| 40 | 40.2 | 1.8 | 1373 | 5 | PCR-US95-03032-3 | Sequence 3, App1 |
| 41 | 40.2 | 1.8 | 1373 | 5 | PCR-US95-03032-3 | Sequence 3, App1 |
| 42 | 39.8 | 1.8 | 969 | 2 | US-08-365-486A-27 | Sequence 27, App1 |
| 43 | 39.8 | 1.8 | 969 | 2 | US-08-880-342-27 | Sequence 27, App1 |
| 44 | 39.8 | 1.8 | 8051 | 2 | US-08-576-626A-2 | Sequence 2, App1 |
| 45 | 39.6 | 1.8 | 724 | 4 | US-08-998-416-674 | Sequence 674, App |

ALIGNMENTS

```
RESULT 1
US-09-182-145-34
; Sequence 34, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gueney, Austin J.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Penitca, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; EARLIER FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 34
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-34
Query Match 1.9%; Score 41.6; DB 4; Length 1212;
Best Local Similarity 73.6%; Pred. No. 0.23;
Matches 53; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
OY 2158 TATCTGCCCAACTAATGATTAAATACAAATCTGTTAAAAA 2217
Db 1119 TATATATAAAAAATAGTAAATAAAAA 1178
OY 2218 AAAAAAAAAA 2229
Db 1179 AAAAAAAAAA 1190
RESULT 2
US-09-182-145-35/C
; Sequence 35, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
```

```

Query Match 1.9%; Score 41.6; DB 4; Length 1212;
Best Local Similarity 73.6%; Pred. No. 0.23;
Matches 53; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 2158 TATAGTCCCAACTAATGAGTTTAAATATACAAATTCGTGTTAAAAAAAAAAAAAAAA 2217
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 94 TATATATAAAAAAAAATTAGTGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 35
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 2218 AAAAAAAAAAAAA 2229
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 34 AAAAAAAAAAAAA 23
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
US-08-086-439C-2
Sequence 2, Application US/08086439C
Patent No. 5468615
GENERAL INFORMATION:
APPLICANT: Chio, Christopher L.
APPLICANT: Huff, Rita M.
TITLE OF INVENTION: A Synthetic Gene for D4 Dopamine
TITLE OF INVENTION: Receptors
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSEE: The Upjohn Company, Corp. Intellectual
STREET: 301 Henrietta Street
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0., Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/086,439C
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Darinley Jr., James D.
REGISTRATION NUMBER: 33,673
REFERENCE/DOCKET NUMBER: 4700
TELECOMMUNICATION INFORMATION:
TELEPHONE: 616-385-5210

```

| | | | | |
|--------------------------|-------|-----------------|-----------|--------------|
| Query Match | 1.8% | Score 41.2; | DB 1; | Length 1161; |
| Best Local Similarity | 60.9% | Pred. No. 0.29; | | |
| Matches 67; Conservative | 0; | Mismatches 43; | Indels 0; | Gaps 0 |

RESULT 4
US-08-434-877-2
: Sequence 2, Application US/08434877

1 GENERAL INFORMATION:
2 APPLICANT: Chio, Christopher L.
3 APPLICANT: Huff, Rita M.
4 TITLE OF INVENTION: A Synthetic Gene for D4 Dopamine
5 TITLE OF INVENTION: Receptors
6 NUMBER OF SEQUENCES: 9
7 CORRESPONDENCE ADDRESSES:
8 ADDRESSEE: The Upjohn Company, Corp. Intellectual
9 ADDRESSEE: Property Law
10 STREET: 301 Henrietta Street
11 CITY: Kalamazoo
12 STATE: Michigan
13 COUNTRY: USA
14 ZIP: 49001
15
16 COMPUTER READABLE FORM:
17 MEDIUM TYPE: Diskette (DS,HD)
18 COMPUTER: Gateway 2000, P5-90
19 OPERATING SYSTEM: PC-DOS/MS-DOS
20 SOFTWARE: Patent In Release #1.0, Version #1.25
21 CURRENT APPLICATION DATA:
22 APPLICATION NUMBER: US/08/434,877
23 FILING DATE: 1 July 1993
24
25 CLASSIFICATION: 435
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Darinley Jr., James D.
28 REGISTRATION NUMBER: 33,673
29 REFERENCE/DOCKET NUMBER: 4700 DVI
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 616-385-5210
32 TELEFAX: 616-385-6897
33 TELEX: 224401
34
35 INFORMATION FOR SEQ. ID NO. 2:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 1161 base pairs
38 TYPE: nucleic acid
39 STRANDEDNESS: single
40 TOPOLOGY: linear
41
42 MOLECULE TYPE: CDNA
43
44 US-08-434-877-2

| | | | | |
|-----------------------|-----------------|-----------------|-----------|--------------|
| Query Match | 1.8%; | Score 41.2; | DB 1; | Length 1161; |
| Best Local Similarity | 60.9%; | Pred. No. 0.29; | | |
| Matches 67; | Conservative 0; | Mismatches 43; | Indels 0; | Gaps 0 |

430 TGGTCTGGGGCTTCGTGCAAGGCAAGACGAGCCGGCTGCTGCGGCGACGCTCA 489

Db 257 TGTGCGCCCTCTCTCTACTCCGAGGTCAGGCGGCGCGCTGCTGAGCCCGCCGCC 316
QY 490 TCCGCTACGCAACCTGGGCAACGCTCATCTCTCGCAGCGCTCA 539
Db 317 TGTGCGAGCGCCCTCATGCGCATGAGCATGCTGTGACCGCCCTCCATC 366

RESULT 5

US-08-475-742-3
Sequence 3, Application US/08475742
Patent No. 6121015

GENERAL INFORMATION:
APPLICANT: O'Malley, Karen L
APPLICANT: Todd, Richard D

TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4 Receptor
FILE REFERENCE: WU 102 CON DIV
CURRENT APPLICATION NUMBER: US/08/475,742
CURRENT FILING DATE: 1995-06-07
EARLIER FILING DATE: 1994-06-16
EARLIER APPLICATION NUMBER: US 08/261,293
EARLIER FILING DATE: 1993-01-28
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3
LENGTH: 1367
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)..(1367)
OTHER INFORMATION: D4 Dopamine Receptor cDNA
PUBLICATION INFORMATION:
AUTHORS: Van Tol, H. H.
AUTHORS: Bunzow, J. R.
TITLE: Cloning of the gene for a human dopamine D4 receptor
TITLE: With high affinity for the antipsychotic clozapine
JOURNAL: Nature
VOLUME: 360
PAGES: 610-614
DATE: 1991

US-08-475-742-3
Query Match 1.8%; Score 41.2; DB 3; Length 1367;
Best Local Similarity 60.9%; Pred. No. 0.32;
Matches 67; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 430 TGTGTCGGGCTCTCTGCAAGGCAAGGAGCGGCGGCTGCTGCGGCGCAGCTCA 489
Db 363 TGTGTCGGGCTCTCTGCTACTCCGAGGTCCAGGTGGCGGCTGCTGAGCCCCGCC 422
QY 490 TCCGCTACGCAACCTGGGCAACGCTCATCTCTGCGCAGCGCTCA 539
Db 423 TGTGCGAGCGCCCTCATGCGCATGAGCATGCTGTGACCGCCCTCCATC 472

RESULT 6

US-08-056-051-1
Sequence 1, Application US/08056051
Patent No. 5516683

GENERAL INFORMATION:
APPLICANT: Grandy, David K
APPLICANT: Bunzow, James R

APPLICANT: Civeill, Olivier
APPLICANT: Van Tol, Hubert H.-M.
TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois

COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/056,051
FILING DATE: 19930429
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: No. 5516683man, Kevin E

REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-C

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234

TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1370 base pairs
TYPE: NUCLEIC ACID

STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: cDNA
FEATURE:

NAME/KEY: 5'UTR
LOCATION: 1..103

FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1268..1370

FEATURE:
NAME/KEY: CDS
LOCATION: 104..1267

US-08-056-051-1
Query Match 1.8%; Score 41.2; DB 1; Length 1370;
Best Local Similarity 60.9%; Pred. No. 0.32;
Matches 67; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 430 TGTGTCGGGCTCTCTGCAAGGCAAGGAGCGGCGGCTGCTGCGGCGCAGCTCA 489
Db 363 TGTGTCGGGCTCTCTGCTACTCCGAGGTCCAGGTGGCGGCTGCTGAGCCCCGCC 422
QY 490 TCCGCTACGCAACCTGGGCAACGCTCATCTCTGCGCAGCGCTCA 539
Db 423 TGTGCGAGCGCCCTCATGCGCATGAGCATGCTGTGACCGCCCTCCATC 472

US-07-928-611-17
Sequence 17, Application US/07928611
Patent No. 5569601
GENERAL INFORMATION:
APPLICANT: Van Tol, Hubert H.M.
APPLICANT: Civeill, Olivier
TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Witcoff, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/928,611

RESULT 9
 US-09-060-694-17
 Sequence 17, Application US/09060694
 Patent No. 620398
 GENERAL INFORMATION:
 APPLICANT: Civeili, Olivier
 APPLICANT: Van Tol, Hubert H.M.
 TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
 STREET: 300 South Wacker Drive
 CITY: Chicago
 STATE: IL
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/060,694
 FILING DATE: 15-Apr-1998
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: NO. 6203998nan, Kevin E
 REGISTRATION NUMBER: 35,303
 REFERENCE/DOCKET NUMBER: 90,1092-MM
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-913-0001
 TELEFAX: 312-913-0002
 TELEX:
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1370 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Wilcoff, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/056,051
FILING DATE: 19930429
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 551663nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1466 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..103
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1364..1466
FEATURE:
NAME/KEY: CDS
LOCATION: 104..1363
US-08-056-051-3
Query Match 1.8%; Score 41.2; DB 1; Length 1466;
Best Local Similarity 60.9%; Pred. No. 0.33;
Matches 67; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 430 TGGTCTCGGCGTCTGTCGAGGCAAGGAGCGAGCGCGCTGCGGCGCAGCTCA 489
DB 363 TGTCTCCGCTCTCTGCTCTACTCTCCGAGGTCCAGGTTGGCGCGTGGCTGCTGAGCCCCCGCC 422
QY 490 TCCGCTAGCCCAACCTGGGCAACGTGCTCATCTCTCGCGCAGCGCTCAGCACC 539
DB 423 TGTGGAGCGCCCTCATGTGGCATGAGACGTGCTGTGTGCAACCGCCCTCCATC 472
RESULT 13
US-07-928-611-19
Sequence 19, Application US/07928611
Patent No. 5569601
GENERAL INFORMATION:
APPLICANT: Van Tol, Hubert H.M.
APPLICANT: Clivell, Olivier
TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Wilcoff, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/928,611
FILING DATE: 19920810
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5569601nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 810-221-8317
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1466 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..103
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1364..1466
FEATURE:
NAME/KEY: CDS
LOCATION: 104..1363
US-07-928-611-19
Query Match 1.8%; Score 41.2; DB 1; Length 1466;
Best Local Similarity 60.9%; Pred. No. 0.33;
Matches 67; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 430 TGGTCTCGGCGTCTGTCGAGGCAAGGAGCGAGCGCGCTGCGGCGCAGCTCA 489
DB 363 TGTCTCCGCTCTCTGCTCTACTCTCCGAGGTCCAGGTTGGCGCGTGGCTGCTGAGCCCCCGCC 422
QY 490 TCCGCTAGCCCAACCTGGGCAACGTGCTCATCTCTCGCGCAGCGCTCAGCACC 539
DB 423 TGTGGAGCGCCCTCATGTGGCATGAGACGTGCTGTGTGCAACCGCCCTCCATC 472
RESULT 14
US-08-487-811A-19
Sequence 19, Application US/08487811A
Patent No. 5883226
GENERAL INFORMATION:
APPLICANT: Clivell, Olivier
APPLICANT: Van Tol, Hubert H.M.
TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,811A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:

NAME: No. 588326nan, Kevin E
 REGISTRATION NUMBER: 35,303
 REFERENCE/DOCKET NUMBER: 90,1092-L
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-913-0001
 TELEFAX: 312-913-0002
 TELEX:
 INFORMATION FOR SEQ ID NO: 19:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1466 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 FEATURE:
 NAME/KEY: 5'UTR
 LOCATION: 1..103
 FEATURE:
 NAME/KEY: 3'UTR
 LOCATION: 1364..1466
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 104..1363
 US-08-487-811A-19

Query Match 1.8%; Score 41.2; DB 2; Length 1466;
 Best Local Similarity 60.9%; Pred. No. 0.33;
 Matches 67; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 430 TGTGTGCGGCTTGTGTCGAGGAGGAGGAGGCGGCTGTGCGGCGGAGCTCA 489
 DB 363 TGTGTCGCGCTTGTGTCGAGGAGGAGGAGGCGGCTGTGCGGCGGAGCTCA 422

OY 490 TCCGCTACGCCAAGCTGCGGCAAGCTGCTCATCTGCGGAGCGTCAGCACC 539
 DB 423 TGTGCGAGCGCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 472

RESULT 15
 US-09-060-694-19
 Sequence 19, Application US/09060694
 Patent No. 6203998
 GENERAL INFORMATION:
 APPLICANT: Civeilli, Olivier
 APPLICANT: Van Tol, Hubert H.M.
 TITLE OF INVENTION: A No. 6203998e1 Human Dopamine Receptor and Uses
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
 STREET: 300 South Wacker Drive
 CITY: Chicago
 STATE: IL
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/060,694
 FILING DATE: 15-APR-1998
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6203998nan, Kevin E
 REGISTRATION NUMBER: 35,303
 REFERENCE/DOCKET NUMBER: 90,1092-MM
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-913-0001
 TELEFAX: 312-913-0002
 TELEX:
 INFORMATION FOR SEQ ID NO: 19:
 SEQUENCE CHARACTERISTICS:

LENGTH: 1466 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 FEATURE:
 NAME/KEY: 5'UTR
 LOCATION: 1..103
 FEATURE:
 NAME/KEY: 3'UTR
 LOCATION: 1364..1466
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 104..1363
 US-09-060-694-19

Query Match 1.8%; Score 41.2; DB 4; Length 1466;
 Best Local Similarity 60.9%; Pred. No. 0.33;
 Matches 67; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 430 TGTGTGCGGCTTGTGTCGAGGAGGAGGAGGCGGCTGTGCGGCGGAGCTCA 489
 DB 363 TGTGTCGCGCTTGTGTCGAGGAGGAGGAGGCGGCTGTGCGGCGGAGCTCA 422

OY 490 TCCGCTACGCCAAGCTGCGGCAAGCTGCTCATCTGCGGAGCGTCAGCACC 539
 DB 423 TGTGCGAGCGCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 472

Search completed: June 8, 2003, 18:16:29
 Job time : 92.8557 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: June 8, 2003, 10:51:03 ; Search time 2950 Seconds
(without alignments)
12237.209 Million cell updates/sec

Title: US-09-622-964-2
Perfect score: 2229
Sequence: 1 cagggagctccaccagccta.....aaaaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|----|--------------------|
| 1 | 1793 | 80.4 | 2453 | 11 | BC015220 Homo sapi |
| 2 | 777.6 | 34.9 | 963 | 14 | B0879880 AGENCOURT |
| 3 | 754.2 | 33.8 | 1699 | 11 | AK006549 Mus muscu |
| 4 | 640.4 | 28.7 | 666 | 13 | B1480798 H2RPE-043 |
| 5 | 582.4 | 26.1 | 592 | 13 | BM663028 UT-E-C10- |
| 6 | 577 | 25.9 | 773 | 10 | BE410951 601303662 |

| | | | | | |
|----|-------|------|------|----|----------|
| 7 | 575 | 25.8 | 585 | 14 | BM707948 |
| 8 | 565.2 | 25.4 | 593 | 10 | BE385296 |
| 9 | 528.8 | 23.7 | 735 | 13 | B1756228 |
| 10 | 525.8 | 23.6 | 537 | 14 | BM691456 |
| 11 | 522 | 22.4 | 522 | 14 | BM707649 |
| 12 | 508 | 22.8 | 526 | 14 | BM685501 |
| 13 | 492.4 | 22.1 | 930 | 14 | B0436824 |
| 14 | 485.2 | 21.8 | 566 | 14 | BM718338 |
| 15 | 475.4 | 21.3 | 503 | 14 | BM685122 |
| 16 | 470.8 | 21.1 | 508 | 14 | BM932117 |
| 17 | 470.8 | 20.9 | 469 | 9 | A1190190 |
| 18 | 464.8 | 20.9 | 469 | 9 | AA307119 |
| 19 | 449.2 | 20.2 | 455 | 12 | BG013943 |
| 20 | 444.6 | 19.9 | 601 | 13 | BG951790 |
| 21 | 443 | 19.9 | 481 | 13 | BM662514 |
| 22 | 437.4 | 19.6 | 531 | 14 | BM685396 |
| 23 | 427.4 | 19.2 | 484 | 14 | N31453 |
| 24 | 405.4 | 18.2 | 419 | 14 | BM718146 |
| 25 | 401.4 | 18.0 | 508 | 10 | BE236901 |
| 26 | 391 | 17.5 | 766 | 10 | BE275846 |
| 27 | 379.8 | 17.0 | 1067 | 13 | BM562042 |
| 28 | 378.4 | 17.0 | 381 | 9 | A1208843 |
| 29 | 365.8 | 16.4 | 371 | 14 | BM694530 |
| 30 | 360.6 | 16.2 | 603 | 9 | AA205892 |
| 31 | 354.8 | 15.9 | 793 | 13 | BG961794 |
| 32 | 347.2 | 15.6 | 444 | 14 | N33227 |
| 33 | 347.2 | 15.3 | 514 | 13 | B1343182 |
| 34 | 307.4 | 13.8 | 599 | 13 | BG951382 |
| 35 | 300 | 13.5 | 525 | 14 | B0345562 |
| 36 | 299.4 | 13.4 | 317 | 14 | N25339 |
| 37 | 293.4 | 13.2 | 305 | 9 | AA318352 |
| 38 | 291.6 | 13.1 | 654 | 10 | BE679896 |
| 39 | 283.4 | 12.7 | 732 | 13 | BG965806 |
| 40 | 282.8 | 12.7 | 297 | 9 | AA991578 |
| 41 | 271.6 | 12.2 | 394 | 13 | BG950641 |
| 42 | 268.8 | 12.1 | 566 | 13 | B1443895 |
| 43 | 246.2 | 11.0 | 271 | 9 | AA317489 |
| 44 | 244.4 | 11.0 | 613 | 9 | AL654606 |
| 45 | 243.4 | 10.9 | 422 | 10 | AM594705 |

ALIGNMENTS

RESULT 1
BC015220 2453 bp mRNA linear HTC 25-JUL-2002
LOCUS Homo sapiens, clone IMAGE:3877806, mRNA.
DEFINITION BC015220
ACCESSION BC015220.1 GI:21955361
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 2453)
Strausberg, R.
Direct Submission
Submitted (01-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT
Contact: MGC help desk
Email: cagabbs-r@mail.nih.gov
Tissue Procurement: DCTP/DP/Genetec
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxll.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN/ at: <http://image.llnl.gov>
 Series: IRK Plate: 14 Row: b Column: 9
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis

This clone has the following problem: incomplete processing.

FEATURES

SOURCE

Location/Qualifiers

1..2453
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3877806"
 /tissue_type="lung, large cell carcinoma"
 /clone_id="NTH_MGC_68"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 BASE COUNT 633 a 711 c 606 g 503 t
 ORIGIN

Query Match 80.4% Score 1793; DB 11; Length 2453;

Best Local Similarity 95.6% Pred. No. 1,le-289;

Matches 1887; Conservative 0; Mismatches 5; Indels 81; Gaps 1;

257 GCTGGCCCTCAGAGAAACAACACCTGATGTTGAGAAACTGACTGTGATTCGACAG 316
 557 GCTGGCCCTCAGAGAAACAACACCTGATGTTGAGAAACTGACTGTGATTCGACAG 616
 317 CTACATCCAGCTCATCCCATTTCTGCTGGGCTTCTACGAGAGCTGCTGCTGAC 376
 617 CTACATCCAGCTCATCCCATTTCTGCTGGGCTTCTACGAGAGCTGCTGCTGAC 676
 377 CCGCTGTGGAACACAGTACAGAACCTGCGTGGCCGACCCGCTCATGAGCTGCTG 436
 677 CCGCTGTGGAACACAGTACAGAACCTGCGTGGCCGACCCGCTCATGAGCTGCTG 736
 437 GGGCTTCGTGAAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 496
 737 GGGCTTCGTGAAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 796
 497 GGGCAACCTGGGCAACAGTCTATCTGCGAGAGCTCAGACAGAGAGAGAGAGAG 556
 797 GGGCAACCTGGGCAACAGTCTATCTGCGAGAGCTCAGACAGAGAGAGAGAGAG 856
 857 CCGCAGCGCCAGACACCTGCTGCAACAGAGCTTTATGACTCCGGCAGAAACAGAG 916
 557 CCGCAGCGCCAGACACCTGCTGCAACAGAGCTTTATGACTCCGGCAGAAACAGAG 616
 857 CCGCAGCGCCAGACACCTGCTGCAACAGAGCTTTATGACTCCGGCAGAAACAGAG 916
 617 GGAAGAACTGAGACCTTCCACACAACTGTTCTGAGGAGCTGAGGAGTGGTTC 676
 917 GGAAGAACTGAGACCTTCCACACAACTGTTCTGAGGAGCTGAGGAGTGGTTC 976
 677 GGAAGAACTGAGACCTTGGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 736
 977 GGAAGAACTGAGACCTGGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1036
 737 GGAAGAACTGAGACCTGGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 796
 1037 GGAAGAACTGAGACCTGGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1096
 797 TATCCCACTGAGTATACAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 856
 1097 TATCCCACTGAGTATACAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1156
 857 TCTAGTGGGCGGAGTTCGAAACAGGCAAGGCTTACCTGAGGAGTGGAGTGGAG 916
 1157 TCTAGTGGGCGGAGTTCGAAACAGGCAAGGCTTACCTGAGGAGTGGAGTGGAG 1216
 917 CGTTGTCGCGGCTTACAGTTCCTGCAAGTTCCTTATGTTGGCTGAGAGTGGAG 976
 1217 CGTTGTCGCGGCTTACAGTTCCTGCAAGTTCCTTATGTTGGCTGAGAGTGGAG 1269

977 AGACAGCTCATCAACCCCTTTGGAGAGATGATGATGATTTGAGCAACAGTGGATTGT 1036
 1270 ----- 1269
 1037 CGACAGAAATTTGAGAGTGTCTCTGTTGGCTGTGATGATGATGATGATGATGAT 1096
 1270 -----AGGTGTCTCTGTTGGCTGTGATGATGATGATGATGATGATGATGAT 1315
 1097 GATGAGCCCGGACATGATACGATGATGATGATGATGATGATGATGATGATGATGAT 1156
 1316 GATGAGCCCGGACATGATACGATGATGATGATGATGATGATGATGATGATGATGAT 1375
 1157 GCGCCAGTTCGCTGAGAGCTCTTATGAGGCTCCACCTTCAACATGACGCTGAAACAA 1216
 1376 GCGCCAGTTCGCTGAGAGCTCTTATGAGGCTCCACCTTCAACATGACGCTGAAACAA 1435
 1217 GGAAGTGAAGTTCACAGCCCAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1276
 1436 GGAAGTGAAGTTCACAGCCCAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1495
 1277 CCGCTTCCTAGAGCCGACAGTCCATGATCAGATCAGATCAGATCAGATCAGATCAG 1336
 1496 CCGCTTCCTAGAGCCGACAGTCCATGATCAGATCAGATCAGATCAGATCAGATCAG 1555
 1337 ACTACTGTGGCCCAAGAGGAGATCCCTTCTCAGAGAGGCTCCCAACAAACACAGAG 1396
 1556 ACTACTGTGGCCCAAGAGGAGATCCCTTCTCAGAGAGGCTCCCAACAAACACAGAG 1615
 1397 AGCCAAACAGACCTTAAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1456
 1616 AGCCAAACAGACCTTAAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1675
 1457 GCGCTTCAAGTGTGGCCCACTGATCAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1516
 1676 GCGCTTCAAGTGTGGCCCACTGATCAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1735
 1517 CCTCAGCCCACTCCCATGTTCTTCCCTAGAGAGATCAGAGAGAGAGAGAGAGAGAG 1576
 1736 CCTCAGCCCACTCCCATGTTCTTCCCTAGAGAGATCAGAGAGAGAGAGAGAGAGAG 1795
 1577 TGTACAGGAGATGAGACCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1636
 1796 TGTACAGGAGATGAGACCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1855
 1637 AGTTTGAATTTGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1696
 1856 AGTTTGAATTTGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1915
 1697 AGTGAG 1756
 1916 AGTGAG 1975
 1757 CCTCAAG 1816
 1976 CCTCAAG 2035
 1817 GGAATCTTATGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1876
 2036 GGAATCTTATGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2095
 1877 GGAATCTTATGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1936
 2096 GGAATCTTATGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2155
 1937 ACAGCCATACAGTGTCCACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1996
 2156 ACAGCCATACAGTGTCCACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2215
 1997 AGCTTATATGATTAATAATCCAGAGTACAGCTTAAATGCTTTATTCATTAATAAC 2056
 2216 AGCTTATATGATTAATAATCCAGAGTACAGCTTAAATGCTTTATTCATTAATAAC 2275
 2057 TGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2116

DB 2276 TGTGAAGGTAGACGACCATTTGAAACATTAACTGACTGTGATTCAGAGTCCGG 2335
OY 2117 AACCTTAGTCTATCTGTAATCCAGACGACCACTAGTATATGTCGCCAACTAAG 2176
DB 2336 AACCTTAGTCTATCTGTAATCCAGACGACCACTAGTATATGTCGCCAACTAAG 2395
OY 2177 AGTTTAATTAATACAAATACCTGTTAAAAA 2229
DB 2396 AGTTTAATTAATACAAATACCTGTTAAAAA 2448
RESULT 2
BQ879880 963 bp mRNA linear EST 16-AUG-2002
LOCUS AGENCOURT 8241531 lupski_dorsal.root.ganglion Homo sapiens cDNA
DEFINITION clone IMAGE:6180559 5', mRNA sequence.
ACCESSION BQ879880.1 GI:22271888
VERSION BQ879880.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13563 row: 1 column: 08
High quality sequence stop: 623.
FEATURES
Source location/Qualifiers
1..963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6180559"
/clone_lib="Lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/note="Vector: PCMV-SPORE6 (Life Technologies); site_1:
NotI; site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGCTCCG-3' and
5'-GACTACTCTAGATCGGAGCGCGCCGCTT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

BASE COUNT 205 a 289 c 265 g 202 t 2 others
ORIGIN
Query Match 34.9%; Score 777.6; DB 14; Length 963;
Best Local Similarity 96.2%; Pred. No. 3.2e-120;
Matches 795; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

OY 445 TCGAAGGACGAGGACGAGGAGCGGCTGCTGCGCGACGCTATCCGCTAGCCCAAC 504
DB 26 TCGAAGGACGAGGACGAGGAGCGGCTGCTGCGCGACGCTATCCGCTAGCCCAAC 85
OY 505 TGGGCAACGTCGTCATCTCCGCGACGCTACGACCGAGTCTACAGCGCTTCCCGACG 564
DB 86 TGGGCAACGTCGTCATCTCCGCGACGCTACGACCGAGTCTACAGCGCTTCCCGACG 145

OY 565 CCCAGCAGCTGCTGACGAGCGCTTTATGACTCCGCGACAAACAAAGACTTGGAGAAC 624
DB 146 CCCAGCAGCTGCTGACGAGCGCTTTATGACTCCGCGACAAACAAAGACTTGGAGAAC 205
OY 625 TGAGGCTACCAACAAATCTTCTGGGTGCGCTGGGTGGTGGTGGCAACCTGTAATGA 684
DB 206 TGAGGCTACCAACAAATCTTCTGGGTGCGCTGGGTGGTGGTGGCAACCTGTAATGA 265
OY 685 AGGCGTGGCTGGAGGTGCAATCCGGAGCCCTATCCCTGCTCCAGACCTGCTGAGAGA 744
DB 266 AGGCGTGGCTGGAGGTGCAATCCGGAGCCCTATCCCTGCTCCAGACCTGCTGAGAGA 325
OY 745 TGAACACCTTGGCTACTGAGTGTGACACCTGTATCCCTACGACTGGATATATCCAC 804
DB 326 TGAACACCTTGGCTACTGAGTGTGACACCTGTATCCCTACGACTGGATATATCCAC 385
OY 805 TGGTGTATACACAGGTGTGACCTGTGGCGGTGTACAGCTTCTTCCGACTGTCTAGTTG 864
DB 386 TGGTGTATACACAGGTGTGACCTGTGGCGGTGTACAGCTTCTTCCGACTGTCTAGTTG 445
OY 865 GGGCGAGTTTCTGAACCCAGCCAGCCCTACCTGAGCTGAGCTGCTGTTGTC 924
DB 446 GGGCGAGTTTCTGAACCCAGCCAGCCCTACCTGAGCTGAGCTGCTGTTGTC 505
OY 925 CGGCTTTCACGTTCTGACAGTTCTTCTTATGTTGGCTGCTGAAGTGGACAGCAGC 984
DB 506 CGGCTTTCACGTTCTGACAGTTCTTCTTATGTTGGCTGCTGAAGTGGACAGCAGC 565
OY 985 TCATCAACCCCTTTGGAGAGATGATGATTTTGGAGCAACCTGGATGTGACAGCA 1044
DB 566 TCATCAACCCCTTTGGAGAGATGATGATTTTGGAGCAACCTGGATGTGACAGCA 625
OY 1045 ATTTCGAGGTCCTGTTGGCTGTTGGCTGTTGGCTGTTGGCTGTTGGCTGTTGGCTG 1104
DB 626 ATTTCGAGGTCCTGTTGGCTGTTGGCTGTTGGCTGTTGGCTGTTGGCTGTTGGCTG 685
OY 1105 CGGACATGTACTGGAATTAAGCCGAGCCAGCCAGCCCTACACAGCTGCTTCCGCCAGT 1164
DB 686 CGGACATGTACTGGAATTAAGCCGAGCCAGCCAGCCCTACACAGCTGCTTCCGCCAGT 745
OY 1165 TCCGCTGAGGCTCTCTTTATGGGCTCCACCTTCAACATCAGCTGGAACAAAGAGATGG 1224
DB 746 TCCGCTGAGGCTCTCTTTATGGGCTCCACCTTCAACATCAGCTGGAACAAAGAGATGG 805
OY 1225 AGTTCAGGCGCAATCAAGAGAGAGAGAGATGCTCAGCTGCGCAT 1270
DB 806 AATTCGAGCGCAATCAAGAGAGAGAGAGATGCTCAGCTGCGCAT 851
RESULT 3
AK006549 1699 bp mRNA linear HTC 19-JAN-2002
LOCUS AK006549
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:1700030H21.vitel1iform macular dystrophy 2 homolog
(human), full insert sequence.
ACCESSION AK006549
VERSION AK006549.1 GI:12839710
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) adult male testis cDNA to mRNA,
clone:1700030H21.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognath; Muridae; Murinae; Mus.
REFERENCE Carninci, P. and Hayashizaki, Y.
AUTHORS High-efficiency full-length cDNA cloning
TITLE Meth. Enzymol. 303, 19-44 (1999)
JOURNAL MEDLINE 99279253
PUBMED 10349636
REFERENCE Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

| QY | 873 | TTTTGGAACCCAGCCAAAGCCCTACCCCTGGCCATGAGCTGGAGACCTGTTGTGGCCGTC | 932 |
|------------|--|--|--------|
| Db | 600 | TTTCTGAAACCCAAACAGAGACTACCCAGGCCATGATGATCTGTGTGTGCTGTCTTC | 659 |
| QY | 933 | ACGTTCTCGACGTTCTTCTCTATGTTGGCTGCTGGAAGGTGGCAGACAGCTATCTAAC | 992 |
| Db | 660 | ACAAATCTCGAATTTCTTATTTCTACATGGCGTGGCTGAAGGTGGCCAAACAGCTCATCAAC | 719 |
| QY | 993 | CCCTTTTGGAGAGGATGATGATTTTATAGACCAACTGGATTGTGACAGCAATTTGCAG | 1055 |
| Db | 720 | CCCTTCGGGGAGGAGATGATGATTTTGAAGCTAACTGATCATTTGACAGAAACCTGCAG | 779 |
| QY | 1053 | GTGTCGCCGTTGGCGTGGTGGATGAGATGTGACACAGGACCTGCCGTGGATGGACCCGACATG | 1112 |
| Db | 780 | GTGTCCCTGTTGTCTCGTGGATGGATGGATGACACAGAACTTGGCTCCCATGGAAAGTATACAT | 839 |
| QY | 1113 | TACTGGAAATTAACCCGAGCCACAGCCCCCTTACACAGCTGTCTCCGCCACATGCTGTGGA | 1172 |
| Db | 840 | TACTGGAAAGAGGAGCGGCTCAGCGCCGCTTACACAGCTGTCTTGGCAGTTCTCGCGCG | 899 |
| QY | 1173 | GCCTCCTTTATGGGCTCCACTTTCACATCAGCTTGAAACAAAGAGAGATGGATTCAG | 1232 |
| Db | 900 | CATTCCTTTCATGGGCTCCACCTTCACATCAGCTTAAAGAAAGAAAGACTTGTAGAGCTTTGG | 959 |
| QY | 1233 | CCCAATCAGGAGG-----ACGAGAGAGATGCTCAACGCTGGCATATTGGCCGC | 1280 |
| Db | 960 | TCAAAAGAGAGGCTGACACAGGATTAAGAAAGAGAGTGGCTTATGACGACACATAGCTGCG | 1019 |
| QY | 1281 | TTCTTAGGCTTCAGCTGCCATGATCACCATCTCTCCAGGGCAAACTCAGGACCAACTA | 1340 |
| Db | 1020 | TTCTTAGGACTGCAACCCAAAACCTACCATCTTCCCTTGAAGAACTTAAAGACCAACTA | 1079 |
| QY | 1341 | CTGTGGCCCAAGAGGGAATCCCTTCTCCAGAGGGGCTGCCCAAAACCAACAGGACGCC | 1400 |
| Db | 1080 | TTGTGTTCTTAAGA-----ATCCCTCTCTCGAAGGCCAGTGTAAAGATGCCAACCCAGAA | 1133 |
| QY | 1401 | AAACAGAACGTTTAAAGGGGCGCAGGAAGACAAACAAGCCCTGGAAGCTTAAAGCTGTGGAGCC | 1460 |
| Db | 1134 | AAACGAAAG-----ATGTGTGGAATTTTAAAGGCTGTGACCTC | 1172 |
| QY | 1461 | TTCAAGTCTGGCCCACTGTATCAGAGCCAGGCTACTACAGTGGCCCCACAGACGCCCTC | 1520 |
| Db | 1173 | TTGAAAGTGTCTTCCAAAGTTTAAAGAGAGAGGCTCCCATTTGTGGCCACAGGACACCCAGC | 1232 |
| QY | 1521 | AGCCCACTCC | 1531 |
| Db | 1233 | AGCCCACTCC | 1243 |
| RESULT 4 | | | |
| BI480798 | | 666 bp | linear |
| LOCUS | | | |
| DEFINITION | H2PE-0430 Human Retinal Pigment Epithelium (2) Homo sapiens cDNA | | |
| | 5' similar to vitelliform macular dystrophy (Best disease, mRNA | | |
| | sequence. | | |
| ACCESSION | BI480798 | | |
| VERSION | BI480798.1 | GI:18998607 | |
| KEYWORDS | EST. | | |
| SOURCE | human. | | |
| ORGANISM | Homo sapiens | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | |
| REFERENCE | 1 (bases 1 to 666) | | |
| AUTHORS | Buraczynska, M., Mears, A.J., Zarepari, S., Farjo, R., Filipova, E., | | |
| | Yuan, Y., Mackie, S.P., Hughes, B. and Swarcop, A. | | |
| TITLE | Towards an expression profile of native human retinal pigment | | |
| | epithelium: Identification of a non-redundant set of more than 1100 | | |
| | genes | | |
| JOURNAL | Unpublished (2001) | | |
| COMMENT | Contact: Swarcop, A. | | |
| | Department of Ophthalmology and Visual Sciences | | |
| | Kellogg Eye Center, University of Michigan | | |

ACCESSION BM663028
 VERSION BM663028.1 GI:18968017
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 592)
 TITLE Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 Seq primer: M13 Forward
 POLYA-yes.

FEATURES
 source
 1..592
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UT-E-C10-aad-h-10-0-UT"
 /clone_lib="UT-E-C10"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site:1: EcoR I; Site:2: Not I; UT-E-C10 is a CDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand CDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded CDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand CDNA contains a library tag sequence that is located between the Not I site and the (GT)18 tail. The sequence tag for this library is ACTTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
 TAG_LIB-UT-E-C10
 TAG_TISSUE-RPE and Choroid
 TAG_SEQ-ACTTA"

BASE COUNT 133 a 111 c 142 g 206 t
 ORIGIN
 Query Match 26.1%; Score 582.4; DB 13; Length 592;
 Best Local Similarity 99.0%; Pred. No. 1.3e-87;
 Matches 586; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1627 GGGCCAGAAAAGTTTGAATTCCTCTCGAGAGCGAGGCGCTTGATGAGACCCAG 1686
 |||
 DB 592 GGGCCAGAAAAGTTTGAATTCCTCTCGAGAGTGATGGGCGCTTGATGAGACCCAG 533
 |||
 QY 1687 AAGATCTCAAGTGAAGGAGGAAAAGTGTGAAGTTAACTGACGAGATGCCAGATCC 1746
 |||
 DB 532 AAGATCTCAAGTGAAGGAGGAAAAGTGTGAAGTTAACTGACGAGATGCCAGATCC 473
 |||
 QY 1747 CCGAAAATCACTCAAGAACCTTTGGAACATACCAACCAACATACACATACACTCA 1806
 |||
 DB 472 CCGAAAATCACTCAAGAACCTTTGGAACATACCAACCAACATACACATACACTCA 413

QY 1807 AAGATCATGATGATCTTATTTGGCCTTGAGAAAACAGGAGTGAACATTTCTTAACCTG 1866
 |||
 DB 412 AAGATCATGATGATCTTATTTGGCCTTGAGAAAACAGGAGTGAACATTTCTTAACCTG 353
 |||
 QY 1867 CTTCCTTAAGGGGAAAGCTTCGCGAGCAGGCTCCTGCTGTGTGTAACACAGAGACAC 1926
 |||
 DB 352 CTTCCTTAAGGGGAAAGCTTCGCGAGCAGGCTCCTGCTGTGTGTAACACAGAGACAC 293
 |||
 QY 1927 TGATCCAGTACAGCCATACAGCTGTCCACACTGGAAGAACGTCCTTAACAGCTGTA 1986
 |||
 DB 292 TGATCCAGTACAGCCATACAGCTGTCCACACTGGAAGAACGTCCTTAACAGCTGTA 233
 |||
 QY 1987 TCAATGCTTACCTTAT 2046
 |||
 DB 232 TCAATGCTTACCTTAT 173
 |||
 QY 2047 TCATATAAACTGTGAAGCTGAGCTGAGACCACTTGAACATTTAACTACAGCTGGATT 2106
 |||
 DB 172 TCATATAAACTGTGAAGCTGAGCTGAGACCACTTGAACATTTAACTACAGCTGGATT 113
 |||
 QY 2107 CAGAGTGGGAAACCTTATGTTCTATCTGATTCAGACAGACCACTTATGATGCTGCC 2166
 |||
 DB 112 CAGAGTGGGAAACCTTATGTTCTATCTGATTCAGACAGACCACTTATGATGCTGCC 53
 |||
 QY 2167 CAACTATGAGTTTAAATTAATACAAATACCTGTTAAAAA 2218
 |||
 DB 52 CAACTATGAGTTTAAATTAATACAAATACCTGTTAAAAA 1

RESULT 6
 BE410951
 LOCUS BE410951 773 bp mRNA linear EST 21-JUL-2000
 DEFINITION 601303662p1 NIH_MGC_21 Homo sapiens CDNA clone IMAGE:3638175 5',
 mRNA sequence.
 BE410951
 BE410951.1 GI:9347401
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 773)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@pds-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
 Plate: LNCM338 row: 1 column: 16
 High quality sequence stop: 662.

FEATURES
 source
 1..773
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3638175"
 /clone_lib="NIH_MGC_21"
 /tissue_type="Choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: placenta; Vector: pOT87; Site:1: XhoI; Site:2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 204 a 239 c 176 g 154 t

QY 2096 GACTGTGATTCAGAGTCGGGAAACCTTATGTTCTATCTATGATCA 2140
 DB 541 GACTGTGATTCAGAGTCGGGAAACCTTATGTTCTATCTATGATCA 585
 RESULT 8
 BE385296 593 bp mRNA linear EST 21-JUL-2000
 LOCUS 601277572F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3618665 5'
 DEFINITION mRNA sequence.
 ACCESSION BE385296
 VERSION BE385296.1 GI:9330661
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 593)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTF
 cDNA Library Preparation: Ling Hong/Rubln Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
 Plate: L1CM287 row: 0 column: 18
 High quality sequence stop: 593.
 Location/Qualifiers
 1..593
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3618665"
 /clone_11b="NIH_MGC_20"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACAG(G). Size selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using Zap-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 176 a 177 c 132 g 108 t
 ORIGIN
 Query Match 25.4%; Score 565.2; DB 10; Length 593;
 Best Local Similarity 99.3%; Pred. No. 9.7e-85;
 Matches 578; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 1268 CATTGATGCGGCTGCTGAGGCTGAGTCCATGATACCATCTCCAGGCAATC 1327
 DB 1 CATTGATGCGGCTGCTGAGGCTGAGTCCATGATACCATCTCCAGGCAATC 60
 QY 1328 AAGGACCAAACTACTGTGGCCCAAGAGGAATCCCTTCCACGAGGGCTGCC-AAAA 1386
 DB 61 AAGGACCAAACTACTGTGGCCCAAGAGGAATCCCTTCCACGAGGGCTGCC-AAAA 120
 QY 1387 ACCACAGGACCAACACAGAGTGTAGGGCCCAAGAGACAAAGGCTGAGAGTTA 1446
 DB 121 ACCACAGGACCAACACAGAGTGTAGGGCCCAAGAGACAAAGGCTGAGAGTTA 180
 QY 1447 AGGCTGTGAGGCTTCAAGTCTGAGCCCACTGTATCAGAGCCAGGCTACTAGTCCC 1506
 DB 181 AGGCTGTGAGGCTTCAAGTCTGAGCCCACTGTATCAGAGCCAGGCTACTAGTCCC 240
 QY 1507 CACAGAGCCCTCAGGCCCACTCCATGTTCTTCCCTTGAACATCAGGCGCTCAA 1566

DB 241 CACAGAGCCCTCAGGCCCACTCCATGTTCTTCCCTTGAACATCAGGCGCTCAA 300
 QY 1567 AGCTTCAAGTGTACAGGCAATGACACCAAGCAAAAGCTTAAAGACTGTGATCTG 1626
 DB 301 AGCTTCAAGTGTACAGGCAATGACACCAAGCAAAAGCTTAAAGACTGTGATCTG 360
 QY 1627 GGGCCAGAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGCTTGTATGAGACACCG 1686
 DB 361 GGGCCAGAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGCTTGTATGAGACACCG 420
 QY 1687 AAGTATCTCAGTGTAGGAGAGAAAGTGTGAGTTTAACTGACGATATGCCAGATCC 1746
 DB 421 AAGTATCTCAGTGTAGGAGAGAAAGTGTGAGTTTAACTGACGATATGCCAGATCC 480
 QY 1747 CCGAAATCAGCTCAAGAAAGCTTGGAAACATACCAACCAATACACTACATCA 1806
 DB 481 CCGAAATCAGCTCAAGAAAGCTTGGAAACATACCAACCAATACACTACATCA 540
 QY 1807 AAGATCAGATGATCCTTATTTGGGCTTGAAAAACAGGATG 1848
 DB 541 AAGATCAGATGATCCTTATTTGGGCTTGAAAAACAGGATG 582
 RESULT 9
 BI756228 735 bp mRNA linear EST 25-SEP-2001
 LOCUS 603024265F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5194649 5'
 DEFINITION mRNA sequence.
 ACCESSION BI756228
 VERSION BI756228.1 GI:15747806
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 735)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: L1AM11487 row: a column: 18
 High quality sequence stop: 577.
 Location/Qualifiers
 1..735
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5194649"
 /clone_11b="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-Sport6; Site:1: NotI;
 Site:2: EcoRV (destroyed). RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH-MGC Library."
 BASE COUNT 135 a 229 c 214 g 157 t
 ORIGIN
 Query Match 23.7%; Score 528.8; DB 13; Length 735;
 Best Local Similarity 92.6%; Pred. No. 1.1e-78;
 Matches 613; Conservative 0; Mismatches 37; Indels 12; Gaps 5;

```

OY 257 GCTGGCCCTCAGGAGAGAACACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 316
DB 59 GCTGGCCCTCAGGAGAGAACACAGCTGATGTTGAGAACTGACTCTGTATGCGACAG 118
OY 317 CTATATCCAGCTCATCCCATTTCTCTGCTGCTGGGCTTCTACGTGACGTGCTGAC 376
DB 119 CTATATCCAGCTCATCCCATTTCTCTGCTGCTGGGCTTCTACGTGACGTGCTGAC 178
OY 377 CCGCTGGTGGAGAACAGTACGAGAACCTTCCCTGGCCCAACCGCTCATGACCTGTGTC 436
DB 179 CCGCTGGTGGAGAACAGTACGAGAACCTTCCCTGGCCCAACCGCTCATGACCTGTGTC 238
OY 437 GGGCTTGTGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 496
DB 239 GGGCTTGTGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 298
OY 497 GCGCAACCTGGGCAACAGTCTCATCTCTGCGAGCGCTGACAGCCGAGCTACAGAGCGCTT 556
DB 299 GCGCAACCTGGGCAACAGTCTCATCTCTGCGAGCGCTGACAGCCGAGCTACAGAGCGCTT 358
OY 557 CCGCAGGCGCCGACAGCTGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 616
DB 359 CCGCAGGCGCCGACAGCTGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 418
OY 617 GGAGAACTGAGCTACACAGACACATGTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
DB 419 GGAGAACTGAGCTACACAGACACATGTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 478
OY 677 GTCATATGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
DB 479 GTCATATGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538
OY 737 GAGCAGATGAGAACACCTT-GCGTACTCAGTGTGAGAACCTGATGCTT-ACGATGAGAT 794
DB 539 GAGCAGATGAGAACACCTT-GCGTACTCAGTGTGAGAACCTGATGCTT-ACGATGAGAT 598
OY 795 AGTATCCCACTGGTGTATACAC-----AGTGTGACTGTGGCGGTGTACAGCTTCTT 847
DB 599 AGTATCCCACTGGTGTATACAC-----AGTGTGACTGTGGCGGTGTACAGCTTCTT 658
OY 848 CCGTACTTGTGAGTGGG--GCGGAGTTTCTGAA-CCGAGCCAGAGGCTTACCTGGGCA 904
DB 659 CCGTACTTGTGAGTGGG--GCGGAGTTTCTGAA-CCGAGCCAGAGGCTTACCTGGGCA 718
OY 905 TG 906
DB 719 TG 720

RESULT 10
LOCUS BM691456 537 bp mRNA linear EST 28-FEB-2002
DEFINITION U1-E-C11-abe-f-10-0-U1.r1 U1-E-C11 Homo sapiens cDNA clone
ACCESSION BM691456
VERSION BM691456.1 GI:19004714
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 537)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

```

```

FEATURES
SOURCE
location/Qualifiers
1..537
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="U1-E-C11-abe-f-10-0-U1"
/clone_1bp="U1-E-C11"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pRT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
U1-E-C11 is a normalized cDNA library containing the
following tissue(s): RPE and Choroid. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pRT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is ACCGA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI)."
BASE COUNT 133 a 172 c 139 g 93 t
Query Match 23.6%; Score 525.8; DB 14; Length 537;
Best Local Similarity 99.6%; Pred. No. 3.7e-78;
Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1020 GAGACCACTGGATGTGACAGAGAAATTGACGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
DB 8 GAGACCACTGGATGTGACAGAGAAATTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 67
OY 1080 CACCGAGACCTGCTCGATGAGAGCCGAGCAATGTAATTAACCCGAGCAGACAGCC 1139
DB 68 CACCGAGACCTGCTCGATGAGAGCCGAGCAATGTAATTAACCCGAGCAGACAGCC 127
OY 1140 CCTACACAGCTGCTTCCGCCAGTTCCTGAGCCCTCTTATGAGGCTCCACCTTCAAC 1199
DB 128 CCTACACAGCTGCTTCCGCCAGTTCCTGAGCCCTCTTATGAGGCTCCACCTTCAAC 187
OY 1200 ATCAGCCTGAGCAAGAGAGAGATGAGTTCACGCCCAATCAGAGAGAGAGAGAGATCT 1259
DB 188 ATCAGCCTGAGCAAGAGAGAGATGAGTTCACGCCCAATCAGAGAGAGAGAGAGATCT 247
OY 1260 CAGCTGAGATGATGGCGGCTTCTAGAGCCGTCAGTCCCATGATCACCATCTCCAGG 1319
DB 248 CAGCTGAGATGATGGCGGCTTCTAGAGCCGTCAGTCCCATGATCACCATCTCCAGG 307
OY 1320 GCAAACTCAAGAGACCAAACTACTGTGGCCCAAGAGAGGAATCCCTTCTCCAGAGGCGCTG 1379
DB 308 GCAAACTCAAGAGACCAAACTACTGTGGCCCAAGAGAGGAATCCCTTCTCCAGAGGCGCTG 367
OY 1380 CCAAAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1439
DB 368 CCAAAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 427
OY 1440 AAGCTTAAGCTGTGAGAGCTTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1499
DB 428 AAGCTTAAGCTGTGAGAGAGCTTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 487
OY 1500 AGTGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1548

```

Email: mssoares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrived by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

Db 488 AGTGCACACAGACGCCCTCAGCCACATGCTTCCCGCTAG 536

|||||

RESULT 11
LOCUS BM707649 522 bp mRNA linear EST 28-FEB-2002
DEFINITION U1-E-CII-afs-o-15-0-UI.r1 U1-E-CII Homo sapiens cDNA clone
U1-E-CII-afs-o-15-0-UI 5', mRNA sequence.
BM707649
BM707649.1 GI:19020907
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 522)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
JOURNAL MEDLINE
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES
source
1..522
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="U1-E-CII-afs-o-15-0-UI"
/clone_1lb="U1-E-CII"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
U1-E-CII is a normalized cDNA library containing the
following tissue(s): RPE and Choroid. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
(dT)18 tail. The sequence tag for this library is ACCTA.
This library was created for the program. Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI)."

BASE COUNT 100 a 182 c 130 g 110 t

ORIGIN

Query Match. 23.4%; Score 522; DB 14; Length 522;
Best Local Similarity 100.0%; Pred. No. 1.6e-77;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

46 CGGACCCACCTGGAACCCACCTGACCCACCTGTCGACGCCACCTGGGCA 105
|||||
1 CGGACCCACCTGGAACCCACCTGACCCACCTGTCGACGCCACCTGGGCA 60
|||||

106 TGACATCATCTACACAGCAAGTGGCTATGCGCCGTTGCTCTTCCCGCTGC 165

Db 61 TGACATCATCTACACAGCAAGTGGCTATGCGCCGTTGCTCTTCCCGCTGC 120
|||||

QY 166 TCGTGTGCTGGCGGCGAGCATCTACAGCTGCTATATGGCGATTCTTAATCTTCG 225
|||||

Db 121 TCGTGTGCTGGCGGCGAGCATCTACAGCTGCTATATGGCGATTCTTAATCTTCG 180
|||||

QY 226 TCGTGTGCTGTACATCATCCGCTTTATTTATATAGCTGGCCCTCAGGAAAGCAACAGCTGA 285
|||||

Db 181 TCGTGTGCTGTACATCATCCGCTTTATTTATATAGCTGGCCCTCAGGAAAGCAACAGCTGA 240
|||||

QY 286 TGTGTAGAAAGTGAATCTGTATTTGGAGAGCTACATCCAGCTATCCCATTTCCCTGC 345
|||||

Db 241 TGTGTAGAAAGTGAATCTGTATTTGGAGAGCTACATCCAGCTATCCCATTTCCCTGC 300
|||||

QY 346 TGTGTGCTGTACATCATCCGCTTTATTTATATAGCTGGCCCTCAGGAAAGCAACAGCTGC 405
|||||

Db 301 TCGTGTGCTGTACATCATCCGCTTTATTTATATAGCTGGCCCTCAGGAAAGCAACAGCTGC 360
|||||

QY 406 CGTGGCGGACCGCCCTCATGAGCTGTGCTGGGCTTCTCGAAGGCAAGAGAGCAAG 465
|||||

Db 361 CGTGGCGGACCGCCCTCATGAGCTGTGCTGGGCTTCTCGAAGGCAAGAGAGCAAG 420
|||||

QY 466 GCCGCTGTGGCGGCGAGCTCATCCGCTACAGCCAACTGGGCAAGCTGCTATCTGC 525
|||||

Db 421 GCCGCTGTGGCGGCGAGCTCATCCGCTACAGCCAACTGGGCAAGCTGCTATCTGC 480
|||||

QY 526 GCAGGCTGACACACCGAGTCTACAAAGCGCTTCCCGAGCGCC 567
|||||

Db 481 GCAGGCTGACACACCGAGTCTACAAAGCGCTTCCCGAGCGCC 522
|||||

RESULT 12
LOCUS BM685501 526 bp mRNA linear EST 27-FEB-2002
DEFINITION U1-E-CIO-aad-h-10-0-UI.r1 U1-E-CIO Homo sapiens cDNA clone
U1-E-CIO-aad-h-10-0-UI 5', mRNA sequence.
BM685501
BM685501.1 GI:18995397
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 526)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
JOURNAL MEDLINE
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES
source
1..526
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="U1-E-CIO-aad-h-10-0-UI"
/clone_1lb="U1-E-CIO"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

RESULT 14
BM718338
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

BM718338 566 bp mRNA linear EST 01-MAR-2002
UI-E-EO1-a3a-j-23-0-UI.1 UI-E-EO1 Homo sapiens cDNA clone
BM718338
UI-E-EO1-a3a-j-23-0-UI 5', mRNA sequence.
BM718338.1 GI:19036537
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 566)
Bonaldio,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1..566
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-EO1-a3a-j-23-0-UI"
/clone_lib="UI-E-EO1"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pTZ19-Pac (Pharmacia) with a
modified polylinker; Site:1; Ecor I; Site:2; Not I;
UI-E-EO1 is a normalized cDNA library containing the
following tissue(s): fetal eye. The library was
constructed according to Bonaldio, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an Ecor I
adaptor, digested with Not I, and cloned directionally
into pTZ19-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
CGGGTATACC. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI)."

BASE COUNT 107 a 182 c 145 g 132 t
ORIGIN

Query Match 21.8%; Score 485.2; DB 14; Length 566;
Best Local Similarity 99.4%; Pred. No. 2.2e-71;
Matches 487; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 CTAGTCCGACGACCTTGTGGGATCATCGGACCCAGCCACTGACCCAAAGC 77
|||||
DB 1 CTAGTCCGACGACCTTGTGGGATCATCGGACCCAGCCACTGACCCAAAGC 60
|||||
QY 78 CCACCGCTGACGACCCATGCGTCATGACCATTAACAAGCAAGTGGCTAT 137
|||||
DB 61 CCACCGCTGACGACCCATGCGTCATGACCATTAACAAGCAAGTGGCTAT 120
|||||
QY 138 GCCCGCTTAGGCTCTCTCCGCTCTGCTGCGGGGAGACATCTACAAGCTG 197

DB 121 GCCCGCTTAGGCTCTCTCCGCTCTGCTGCTGCGGGGAGACATCTACAAGCTG 180
|||||
QY 198 CTATATGGCGAGTTCTTAATCTTCTGCTGCTACTACATCATCGGCTTTTATAG 257
|||||
DB 181 CTATATGGCGAGTTCTTAATCTTCTGCTGCTACTACATCATCGGCTTTTATAG 240
|||||
QY 258 CTGGCCCTCAGGAGAACACAGAGCTGATGTTTGAAGTACTCTGATTTGACACAGC 317
|||||
DB 241 CTGGCCCTCAGGAGAACACAGAGCTGATGTTTGAAGTACTCTGATTTGACACAGC 300
|||||
QY 318 TACATTCAGCTCATCCCATTTCTTCTGCTGCTGCTTCTACGTACGCTGCTGACC 377
|||||
DB 301 TACATTCAGCTCATCCCATTTCTTCTGCTGCTGCTTCTACGTACGCTGCTGACC 360
|||||
QY 378 CGCTGCTGGAACAGTACGACAACTGCGCGCCGACCGGCTCATGAGCCTGGTCTG 437
|||||
DB 361 CGCTGCTGGAACAGTACGACAACTGCGCGCCGACCGGCTCATGAGCCTGGTCTG 420
|||||
QY 438 GCGTTCGTCGAAGCAGACGACGACGCGGCTGCTGCGGCGACGCTCATCCGCTTAC 497
|||||
DB 421 GCGTTCGTCGAAGCAGACGACGACGCGGCTGCTGCGGCGACGCTCATCCGCTTAC 480
|||||
QY 498 GCCAACCCTGG 507
|||||
DB 481 GCCAACCCTGG 490
|||||

RESULT 15
BM685122/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

BM685122 503 bp mRNA linear EST 27-FEB-2002
UI-E-EO1-a3l-1-09-0-UI.s1 UI-E-EO1 Homo sapiens cDNA clone
BM685122
UI-E-EO1-a3l-1-09-0-UI 3', mRNA sequence.
BM685122.1 GI:18995018
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 503)
Bonaldio,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Forward
POLYA=Yes.
Location/Qualifiers
1..503
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-EO1-a3l-1-09-0-UI"
/clone_lib="UI-E-EO1"
/tissue_type="fetal eye"
/tissue_type="fetal eye"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pTZ19-Pac (Pharmacia) with a

modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-EJ1 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-PAC vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)₁₈ tail. The sequence tags for this library are: fetal eyes, AGAATCAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
TAG_LIB-UI-E-EJ1
TAG_TISSUE-RPE and Choroid
TAG_SEO-ACCTA"

BASE COUNT 96 a 107 c 123 g 177 t
ORIGIN

Query Match 21.3%; Score 475.4; DB 14; Length 503;
Best Local Similarity 99.8%; Pred. No. 9,7e-70;
Matches 476; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1547 AGAACCATCAGCGCCGCTCAAAAGCTTCACAGTGTACAGGCTATGACACCAAGACAAAG 1606
|||||
Db 494 AGAACCATCAGCGCCGCTCAAAAGCTTCACAGTGTACAGGCTATGACACCAAGACAAAG 435
|||||
QY 1607 CTTAAAGCTGTGAGTTCTGGGGCCCAAGAAAGTTTGATGTCTCTCAGAGAGCGATGG 1666
|||||
Db 434 CTTAAAGCTGTGAGTTCTGGGGCCCAAGAAAGTTTGATGTCTCTCAGAGAGCGATGG 375
|||||
QY 1667 GGCCTTGATGAGACACCCAGAAATCTCAAGTGTAGAGAGAAACCTTGAGAGTTTAACCT 1726
|||||
Db 374 GGCCTTGATGAGACACCCAGAAATCTCAAGTGTAGAGAGAAACCTTGAGAGTTTAACCT 315
|||||
QY 1727 GACGGATATGCCAGAGATCCCGAAATCACCCTCAAGAACCTTGGAACAATCACCAAC 1786
|||||
Db 314 GACGGATATGCCAGAGATCCCGAAATCACCCTCAAGAACCTTGGAACAATCACCAAC 255
|||||
QY 1787 CAACATACACACTACACCAAGATCATGATGATCTTATGCGCTTGGAACAAGAGGA 1846
|||||
Db 254 CAACATACACACTACACCAAGATCATGATGATCTTATGCGCTTGGAACAAGAGGA 195
|||||
QY 1847 TGAAGCACATTCCTAACCTGCTCCTATGAGGATGCTTCGCCAGCAGGCTCCTCACCCTG 1906
|||||
Db 194 TGAAGCACATTCCTAACCTGCTCCTATGAGGATGCTTCGCCAGCAGGCTCCTCACCCTG 135
|||||
QY 1907 TGTGTACACAGCAGGACACTGATCCAGTCAAGCCATACAGCTGTCCACACTGAAGAAC 1966
|||||
Db 134 TGTGTACACAGCAGGACACTGATCCAGTCAAGCCATACAGCTGTCCACACTGAAGAAC 75
|||||
QY 1967 GTGTCTACACAGCAGCTGATCAAAATGCTTATGATTAATGATTAATCCAGACTA 2023
|||||
Db 74 ATGTCTACACAGCAGCTGATCAAAATGCTTATGATTAATGATTAATCCAGACTA 18
|||||

Search completed: June 9, 2003, 03:30:02
Job time : 2956 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 09:30:48 ; Search time 5619.81 Seconds

(without alignments)
11543.117 Million cell updates/sec

Title: US-09-622-964-2

Perfect score: 2229

Sequence: 1 cagggagctcccccagcagccta.....aaaaaaaaaaaaaaaaaaaaa 2229

Scoring table:

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : GenBank1.*

```
1: gb_da.*
2: gb_hg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pa.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_or.*
21: em_ov.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_hg_hum.*
31: em_hg_inv.*
32: em_hg_other.*
33: em_hg_mus.*
34: em_hg_pln.*
35: em_hg_rnd.*
36: em_hg_man.*
37: em_hg_vtl.*
38: em_sy.*
39: em_hgo_hum.*
40: em_hgo_mus.*
41: em_hgo_other.*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|--------|-------|--------|----|-----------|---------------------|
| 1 | 2208.4 | 99.1 | 2210 | 9 | AF057169 | AF057169 Homo sapi |
| 2 | 2166.2 | 97.2 | 2171 | 9 | AF073501 | AF073501 Homo sapi |
| 3 | 2002.4 | 89.8 | 2420 | 9 | AF057170 | AF057170 Homo sapi |
| 4 | 861.4 | 38.6 | 1289 | 9 | AY064707 | AY064707 Sus scrofa |
| 5 | 642.2 | 28.8 | 142092 | 2 | AF139813 | AF139813 Homo sapi |
| 6 | 642.2 | 28.8 | 163915 | 2 | AC087451 | AC087451 Homo sapi |
| 7 | 642.2 | 28.8 | 166867 | 9 | AP003733 | AP003733 Homo sapi |
| 8 | 642.2 | 28.8 | 196080 | 9 | AC004328 | AC004328 Homo sapi |
| 9 | 640.6 | 28.7 | 706 | 9 | HSVMD2P10 | HSVMD2P10 |
| 10 | 640.6 | 28.7 | 112309 | 9 | AC003025 | AC003025 |
| 11 | 640.6 | 28.7 | 133683 | 2 | AC084857 | AC084857 Homo sapi |
| 12 | 576.8 | 25.9 | 1908 | 9 | AF440756 | AF440756 Homo sapi |
| 13 | 552.8 | 24.8 | 160169 | 2 | AC051664 | AC051664 Homo sapi |
| 14 | 538.4 | 24.2 | 1956 | 10 | BC031186 | BC031186 Homo sapi |
| 15 | 538.4 | 24.2 | 1957 | 10 | BC019528 | BC019528 Mus muscu |
| 16 | 535 | 24.0 | 1861 | 9 | AK000139 | AK000139 Homo sapi |
| 17 | 459.2 | 20.6 | 2500 | 9 | AK096459 | AK096459 Homo sapi |
| 18 | 457.6 | 20.5 | 1506 | 9 | AF440758 | AF440758 Homo sapi |
| 19 | 446.2 | 20.0 | 2045 | 9 | AF440757 | AF440757 Homo sapi |
| 20 | 358.4 | 16.1 | 492 | 9 | HSVMD2P11 | HSVMD2P11 |
| 21 | 358.4 | 16.1 | 106648 | 2 | AC004588 | AC004588 |
| 22 | 314.8 | 14.1 | 2874 | 3 | AF218817 | AF218817 Homo sapi |
| 23 | 314.8 | 14.1 | 5610 | 3 | AY061546 | AY061546 Drosophila |
| 24 | 240.8 | 10.8 | 413 | 9 | HSVMD2P04 | HSVMD2P04 |
| 25 | 240.8 | 10.8 | 1666 | 9 | AF052095 | AF052095 Homo sapi |
| 26 | 218.6 | 9.8 | 32022 | 2 | AC019521 | AC019521 Drosophila |
| 27 | 218.6 | 9.8 | 170675 | 3 | AC008139 | AC008139 Drosophila |
| 28 | 218.6 | 9.8 | 192055 | 3 | AC009183 | AC009183 Drosophila |
| 29 | 218.6 | 9.8 | 221888 | 3 | AE003686 | AE003686 Drosophila |
| 30 | 191.6 | 8.6 | 246 | 10 | AF057171 | AF057171 Mus muscu |
| 31 | 189 | 8.5 | 160169 | 2 | AC051664 | AC051664 Homo sapi |
| 32 | 179.4 | 8.0 | 125804 | 2 | AC019853 | AC019853 Drosophila |
| 33 | 179.4 | 8.0 | 163514 | 3 | AC091227 | AC091227 Drosophila |
| 34 | 179.4 | 8.0 | 274351 | 3 | AE003531 | AE003531 Drosophila |
| 35 | 179.4 | 8.0 | 274351 | 3 | G05922 | G05922 human STS W |
| 36 | 175 | 7.9 | 438 | 11 | G05922 | G05922 human STS W |
| 37 | 175 | 7.9 | 1119 | 9 | HUMERRH | HUMERRH |
| 38 | 175 | 7.9 | 1198 | 6 | AX409527 | AX409527 Sequence |
| 39 | 175 | 7.9 | 1198 | 6 | AX409527 | AX409527 Sequence |
| 40 | 161.4 | 7.2 | 239 | 9 | HSVMD2P05 | HSVMD2P05 |
| 41 | 160.4 | 7.2 | 295 | 9 | HSVMD2P07 | HSVMD2P07 |
| 42 | 155 | 7.0 | 289 | 9 | HSVMD2P09 | HSVMD2P09 |
| 43 | 139.8 | 6.3 | 125804 | 2 | AC019853 | AC019853 Drosophila |
| 44 | 139.8 | 6.3 | 163514 | 3 | AC091227 | AC091227 Drosophila |
| 45 | 139.8 | 6.3 | 274351 | 3 | AE003531 | AE003531 Drosophila |

ALIGNMENTS

RESULT 1
AF057169
LOCUS Homo sapiens bestrophin (VMD2) mRNA, alternatively spliced product, complete cds.
ACCESSION AF057169
VERSION AF057169.1 GI:3335158
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 2210)
Petrukhin, K., Kolst, M.J., Bakall, B., Li, W., Xie, G., Markneil, T., Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, M.,


```

Db      1501 GTGCCCCACAGAGCCCCCTCAGGCCACTCCAGTCTTCCCTTGAACCATCAGCGC 1560
Oy      1561 CGTCAAGGCTTCACAGTGTACAGGATAGACACCAAGAACAAAGCTTAAAGACTGTA 1620
Db      1561 CGTCAAGGCTTCACAGTGTACAGGATAGACACCAAGAACAAAGCTTAAAGACTGTA 1620
Oy      1621 GTCTGTGGGCGGCAAGAAAGTTTGAANTGCTTCAGAGAGCGATGGGCGCTTGATGAGC 1680
Db      1621 GTCTGTGGGCGGCAAGAAAGTTTGAANTGCTTCAGAGAGCGATGGGCGCTTGATGAGC 1680
Oy      1681 ACCAGAGATATCTCAAGTAGAGAGAAACCTGTGAGTTAACTGACGATATGCGAG 1740
Db      1681 ACCAGAGATATCTCAAGTAGAGAGAAACCTGTGAGTTAACTGACGATATGCGAG 1740
Oy      1741 AGATCCCCGAAATTCACCTCAAGAAACCTTTGAGACATCCACCAACATACACACTA 1800
Db      1741 AGATCCCCGAAATTCACCTCAAGAAACCTTTGAGACATCCACCAACATACACACTA 1800
Oy      1801 CACTCAAGATACATGATGATCTTATTTGGGCTTTGAAAACAGGATGAAAGCATTCCT 1860
Db      1801 CACTCAAGATACATGATGATCTTATTTGGGCTTTGAAAACAGGATGAAAGCATTCCT 1860
Oy      1861 AACCTGCTCTTAAGGGGATGCTTCGCCAGCCAGGCTCTCACTGTGTGTACACAGCA 1920
Db      1861 AACCTGCTCTTAAGGGGATGCTTCGCCAGCCAGGCTCTCACTGTGTGTACACAGCA 1920
Oy      1921 GGACACTGATCCAGTACAGGATACAGTGTCCACACAGCAAGACGTGCTTACACAG 1980
Db      1921 GGACACTGATCCAGTACAGGATACAGTGTCCACACAGCAAGACGTGCTTACACAG 1980
Oy      1981 CCTGAATCAATGTTAGCTTAATAGATTAATCAATCCAGACTACTGACCTTTAATGCC 2040
Db      1981 CCTGAATCAATGTTAGCTTAATAGATTAATCAATCCAGACTACTGACCTTTAATGCC 2040
Oy      2041 TTTTATATATAAACTGTGAAGCTAGACTGACCAATGGAATTAATTAATCTGAGATC 2100
Db      2041 TTTTATATATAAACTGTGAAGCTAGACTGACCAATGGAATTAATTAATCTGAGATC 2100
Oy      2101 TGGATTCAGAGTCCGGAACCTTATGTTCTATCTGATGATCAAGACGACCACTTATAT 2160
Db      2101 TGGATTCAGAGTCCGGAACCTTATGTTCTATCTGATGATCAAGACGACCACTTATAT 2160
Oy      2161 ACTGCCCAACTAATGATTAATTAATCAAAATCTGTTAAAAAAA 2210
Db      2161 ACTGCCCAACTAATGATTAATTAATCAAAATCTGTTAAAAAAA 2210

RESULT 2
AF073501 2171 bp mRNA linear PRI 02-SEP-1998
LOCUS Homo sapiens vitelliform macular dystrophy protein (VMD2) mRNA,
DEFINITION complete cds.
ACCESSION AF073501
VERSION AF073501.1 GI:3511241
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 2171)
AUTHORS Strohr, H., Marguardt, A., Rivera, A., Cooper, P. R., Nowak, N. J.,
TITLES Shows, T. B., Gerhard, D. S., and Weber, B. H.,
A gene map of the Best's vitelliform macular dystrophy region in
JOURNAL chromosome Res. 8 (1), 48-56 (1998)
MEDLINE 98112782
PUBMED 9445487
REFERENCE 2 (bases 1 to 2171)
AUTHORS Marguardt, A., Strohr, H., Passmore, L., Kraemer, F., Rivera, A., and
TITLES Weber, B. H. F.,
Direct Submission
JOURNAL Submitted (22-JUN-1998) Human Genetics, University, Biozentrum, Am

```

```

FEATURES             Hubland, Wuerzburg 97074, Germany
SOURCE               1. 2171
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="11"
                    /map="11q12-q13.1"
                    /tissue-type="retinal pigment epithelium"
gene                 1. 2171
                    /gene="VMD2"
                    /gene="VMD2"
                    /note="Best disease"
                    /codon_start=1
                    /product="vitelliform macular dystrophy protein"
                    /protein_id="AAC3766.1"
                    /db_xref="GI:3511242"
                    /translation="MTITYSQVANNRLGFSRLLLCMRGSIYKLLYGEFLFLCY
                    IIRFYRLALTEBOIMEKTLKLYDSYIOLIPISVLGFTYLVYTRMMNOYENLPM
                    PRLMSIVSGFEKDEOGRLLRLIRVANGNVILRSVTAAYKRRPSAOHLVOA
                    GEMTFAEHKOLEKLSLPHNMFVWPVWPANLSMKATGSRIRDPILLGSLNEMTLR
                    TCGHLVAYDWTLSILVYOVYTVAVYSFFLTCYGRQPLNPAKATPGHELDLVYVF
                    TELQREFYVGMKVAIBOLINPREDDEETNWIYDRNLQVSLAVDEMHQDLPRMEP
                    DMYNKPEQPPYTAASQFRASFWGTFNLSLKNKEEMEPQNOEDEDAAGILGR
                    FLGLSHDHHPRANRSRKLMPKRESLHLELPRNKKAAKONVROEDNKKAMKLKAV
                    DAFKSAPLYORPGYSAPQTPSPMPFPLPSPSKLSVYGIIDTKSLKTVSSG
                    AKSFELSESDGALMEHPEVQVRKRTYEVFLTDMPELPEHNLKEPLEDSPTNIHTT
                    LKDHMDPWALENREDAHS"
BASE COUNT          540 a 638 c 523 g 470 t
ORIGIN
Query Match          97.2%; Score 2166.2; DB 9; Length 2171;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy      17 CCTATGCCGAGACTTCTGTGGGATCATCGGACCCACCTGGAAACCCACCTGACCCAAG 76
Db      1 CCTATGCCGAGACTTCTGTGGGATCATCGGACCCACCTGGAAACCCACCTGACCCAAG 60
Oy      77 CCCACTGCTCAGGCCACTGCTGGGCGCATGACCATCACTACACAAAGGAGGCTTAA 136
Db      61 CCCACTGCTCAGGCCACTGCTGGGCGCATGACCATCACTACACAAAGGAGGCTTAA 120
Oy      137 TGCCTGTTAGGCTCTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 196
Db      121 TGCCTGTTAGGCTCTCTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Oy      197 GCTATAGGGAGGCTTAACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 256
Db      181 GCTATATGGGAGGTTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Oy      257 GCTGCGCCCTCAGGAGAACACAGCTGATGTTGAGAAATGACTGTTTTCGACAG 316
Db      241 GCTGCGCCCTCAGGAGAACACAGCTGATGTTGAGAAATGACTGTTTTCGACAG 300
Oy      317 CTACATCCAGCTATCCCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 376
Db      301 CTACATCCAGCTATCCCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Oy      377 CCGGTGTGGAACAGAGAGAAACCTGCGCTGGCCGACCGCTCATGAGCTGTGTC 436
Db      361 CCGGTGTGGAACAGAGAGAAACCTGCGCTGGCCGACCGCTCATGAGCTGTGTC 420
Oy      437 GGGCTTCGTGGAAGGACAGAGAGAGCGGCTGCTGCGGCGACGCTCATGCGCTA 496
Db      421 GGGCTTCGTGGAAGGACAGAGAGAGCGGCTGCTGCGGCGACGCTCATGCGCTA 480
Oy      497 CGGCAACCTGGGGAACGCTGCTATCTGCGCAGCTGACAGCCGAGTCTACAAAGCGTT 556
Db      481 CGGCAACCTGGGGAACGCTGCTATCTGCGCAGCTGACAGCCGAGTCTACAAAGCGTT 540
Oy      557 CCCGACGCGCCAGCACTGTGCAACGAGCTTATGACTCCGCGAGAACACAGAGATT 616

```

Db 541 CCCACGCCCGCCAGACCTGTGTCAGACAGGCTTATGACTCCGGGCGAAGACAGACT 600
 Oy 617 GGAGAACTGAGCCCTACCAACACATATGTTCTGGGGTCCCTGGGTGTGTTGGCAACT 676
 Db 601 GGAGAACTGAGCCCTACCAACACATATGTTCTGGGGTCCCTGGGTGTGTTGGCAACT 660
 Oy 677 GTCAATGAAAGCGGTGGAGTGTGAAATCCGGGACCTATCTGCTCCAGAGCGCTGT 736
 Db 661 GTCAATGAAAGCGGTGGAGTGTGAAATCCGGGACCTATCTGCTCCAGAGCGCTGT 720
 Oy 737 GAACAGATGAACACCTTGCCTACAGTGTGAGACACCTGTATGCTACAGTGAATTAG 796
 Db 721 GAACAGATGAACACCTTGCCTACAGTGTGAGACACCTGTATGCTACAGTGAATTAG 780
 Oy 797 TATCCACTGGGTATACAGAGTGTGATCTGCTGCGGTGTACACTTCTTCTGACTTG 856
 Db 781 TATCCACTGGGTATACAGAGTGTGATCTGCTGCGGTGTACACTTCTTCTGACTTG 840
 Oy 857 TCTAGTTGGGGGCGAGTTTCTGMAACCCAGCCAGGCTACCTGGCCATGAGCTGAGCT 916
 Db 841 TCTAGTTGGGGGCGAGTTTCTGMAACCCAGCCAGGCTACCTGGCCATGAGCTGAGCT 900
 Oy 917 CGTGTGCCCTCTTACAGTGTCTGACGTTCTTCTTATGTTGGCTGGCTGAAGTGGC 976
 Db 901 CGTGTGCCCTCTTACAGTGTCTGACGTTCTTCTTATGTTGGCTGGCTGAAGTGGC 960
 Oy 977 AGAGAGCTCATCAACCCCTTGGAGAGATGATATTTTGGAGCAACTGATGTGT 1036
 Db 961 AGAGAGCTCATCAACCCCTTGGAGAGATGATATTTTGGAGCAACTGATGTGT 1020
 Oy 1037 CGACAGGAATTTGCAAGTGTCTGCTGTTGGCTGTGATGATGACACAGACCTGCTCG 1096
 Db 1021 CGACAGGAATTTGCAAGTGTCTGCTGTTGGCTGTGATGATGACACAGACCTGCTCG 1080
 Oy 1097 GATGAGCGCGGACATGCTGGAATTAAGCCGAGCCAGACAGCCCTTACAGCTGCTTC 1156
 Db 1081 GATGAGCGCGGACATGCTGGAATTAAGCCGAGCCAGACAGCCCTTACAGCTGCTTC 1140
 Oy 1157 CGCCAGTTCCGTGAGAGCTCTTATGAGCTCCACCTTCAACATCAGCTGAACAAAGA 1216
 Db 1141 CGCCAGTTCCGTGAGAGCTCTTATGAGCTCCACCTTCAACATCAGCTGAACAAAGA 1200
 Oy 1217 GGAATGAGATTTCAAGCCCAATCAGAGAGACGAGAGAGATGCTCAAGCTGATATTGG 1276
 Db 1201 GGAATGAGATTTCAAGCCCAATCAGAGAGACGAGAGAGATGCTCAAGCTGATATTGG 1260
 Oy 1277 CGGCTTCTAGGCGTGGACCTGATCCCATGATCCATCTCCAGGCGCAACTAAGGACCA 1336
 Db 1261 CGGCTTCTAGGCGTGGACCTGATCCCATGATCCATCTCCAGGCGCAACTAAGGACCA 1320
 Oy 1337 ACTACTGTGGCCCAAGAGGGAATCCCTTCTCAGAGAGGCGCTGCCCAAAAACCAAGGC 1396
 Db 1321 ACTACTGTGGCCCAAGAGGGAATCCCTTCTCAGAGAGGCGCTGCCCAAAAACCAAGGC 1380
 Oy 1397 AGCCAAACAGAACTTAAAGGCGCAGAAAGACAAAGGCGCTGAAGCTTAAAGCTGTGGA 1456
 Db 1381 AGCCAAACAGAACTTAAAGGCGCAGAAAGACAAAGGCGCTGAAGCTTAAAGCTGTGGA 1440
 Oy 1457 CGGCTTCTAGGCGTGGACCTGATCCCATGATCCATCTCCAGGCGCAACTAAGGACCA 1516
 Db 1441 CGGCTTCTAGGCGTGGACCTGATCCCATGATCCATCTCCAGGCGCAACTAAGGACCA 1500
 Oy 1517 CGTACAGCCCACTCCCATGTTCTTCCCTTGAAGACATCAAGGCGCTCAAGCTTCAAG 1576
 Db 1501 CGTACAGCCCACTCCCATGTTCTTCCCTTGAAGACATCAAGGCGCTCAAGCTTCAAG 1560
 Oy 1577 TGTACAGGCGATAGACACCAAGCAAAAGCTTAAAGCTGTGAGTTCTGGGCGCAAGAA 1636
 Db 1561 TGTACAGGCGATAGACACCAAGCAAAAGCTTAAAGCTGTGAGTTCTGGGCGCAAGAA 1620
 Oy 1637 AAGTTTGAATGCTCTCAGAGAGCGATGGGCGCTGATGAGAGACCAAGAAATATCTCA 1696
 Db 1621 AAGTTTGAATGCTCTCAGAGAGCGATGGGCGCTGATGAGAGACCAAGAAATATCTCA 1680

Oy 1697 ACTGAGAGAAAGCTGTGAGTTAACTGACGATATATCCAGAGATCCCGAAATCA 1756
 Db 1681 ACTGAGAGAAAGCTGTGAGTTAACTGACGATATATCCAGAGATCCCGAAATCA 1740
 Oy 1757 CCTCAAGAACTTTGGAACATCAACCAACATACACTATCACTTAAAGATCACT 1816
 Db 1741 CCTCAAGAACTTTGGAACATCAACCAACATACACTATCACTTAAAGATCACT 1800
 Oy 1817 GGATCTTATTTGGGCTTTGGAAGAAACAGGATGAAACACATCTTAACTGTTCTTATG 1876
 Db 1801 GGATCTTATTTGGGCTTTGGAAGAAACAGGATGAAACACATCTTAACTGTTCTTATG 1860
 Oy 1877 GGGATGCTTGGCCAGCCAGGCTCTCACTGTGTGACACACAGACAGATCCAGTC 1936
 Db 1861 GGGATGCTTGGCCAGCCAGGCTCTCACTGTGTGACACACAGACAGATCCAGTC 1920
 Oy 1937 ACAGCCATACAGCTGTCCAGACTGAAAGAAAGCTGTCTCAACAGCTGAATGATGT 1996
 Db 1921 ACAGCCATACAGCTGTCCAGACTGAAAGAAAGCTGTCTCAACAGCTGAATGATGT 1980
 Oy 1997 AGCTTAATAGATTAATAATCCAGACTACTTCACTTAACTGCTTTATTCATTAATCA 2056
 Db 1981 AGCTTAATAGATTAATAATCCAGACTACTTCACTTAACTGCTTTATTCATTAATCA 2040
 Oy 2057 TGTGAACCTGACTGAACCATGGAACATTTAATCACTGACTGTGATTCAGATCGGG 2116
 Db 2041 TGTGAACCTGACTGAACCATGGAACATTTAATCACTGACTGTGATTCAGATCGGG 2100
 Oy 2117 AACCTTATGTTCTATCTGAATCCAGAGACAGCCAGCTTATGATCTGCGCAACTATG 2176
 Db 2101 AACCTTATGTTCTATCTGAATCCAGAGACAGCCAGCTTATGATCTGCGCAACTATG 2160
 Oy 2177 AGTTAATATA 2187
 Db 2161 AGTTAATATA 2171

RESULT 3
 AF057170
 LOCUS
 DEFINITION Homo sapiens bestraphin (VMD2) mRNA, alternatively spliced product,
 complete cds.
 AF057170
 ACCESSION
 VERSION AF057170.1 GI:3335160
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Petrunkhin, K., Koist, M. J., Bakall, B., Li, W., Xie, G., Marknell, T.,
 Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vuojic, M.,
 Bergen, A. A., McGarity-Dugan, V., Figueroa, D., Austin, C. P.,
 Metzker, M. L., Caskey, C. T. and Madellus, C.
 Identification of the gene responsible for Best macular dystrophy
 Nat. Genet. 19 (3), 241-247 (1998)
 JOURNAL JOURNAL
 MEDLINE 98324772
 PUBMED 9662395
 REFERENCE
 AUTHORS Petrunkhin, K.
 TITLE Submitted (03-APR-1998) Human Genetics, Merck Research
 Laboratories, West Point, PA 19486, USA
 JOURNAL Location/Qualifiers
 FEATURES
 source
 1..2420
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q13"
 1..2420
 /gene="VMD2"
 105..1412
 CDS

```
/gene="VMD2"  
/note="Best's macular dystrophy gene; BMD; alternatively  
splice isoform; contains exon 7"  
/codon_start=1  
/product="bestrophin"  
/protein_id="AAC64344.1"  
/db_xref="GI:3335161"  
/translation="MTITTSOVANMARLSSPRLLCWRSGSIYKLYLGEPLFLCY  
IIRFIYALTEEOQLMEKTLTXCDYIOLIPISPLGFTVLVTRMNOYENLPW  
PRLMSLVSGFKDQKQRLIRTLIRINLGNVLILRSVSTAVYRFPASGHLVQA  
GEMTPRHKOLEKLSPHNFVPMWFANLSKAMGRIKIDPILLOSILNENLTA  
TQCGHLTAIDMISIPLYTQVAVYSPFLCLVGRFLNPKAYGHELDIVPVF  
TLDFFFYGMILKVLGRHGRHSGQOLLETRMOCERKYSRVSSQAMPT  
PVIPTAREAGESLIEGRRLMWOSSSLPIERMMLILPGLSTICRCFCLMWR  
CIRTCGMSRSTCGISPSHPROLPPRSVPEPLNAPSTSA"  
ORIGIN  
BASE COUNT      605 a      695 c      612 g      508 t  
  
Query Match      89.8%; Score 2002.4; DB 9; Length 2420;  
Best Local Similarity 91.6%; Pred. No. 0;  
Matches 2216; Conservative 0; Mismatches 1; Indels 203; Gaps 1;  
  
QY      1 CAGGGAGTCCCAACCAAGCTAGTGGCCAGACCTTCTGGGATCATGGACCACCTGGAA 60  
Db      1 CAGGGAGTCCCAACCAAGCTAGTGGCCAGACCTTCTGGGATCATGGACCACCTGGAA 60  
QY      61 CCCCACTGACCCCAACCAAGCTAGTGGCCAGACCTTCTGGGATCATGGACCACCTGGAA 120  
Db      61 CCCCACTGACCCCAACCAAGCTAGTGGCCAGACCTTCTGGGATCATGGACCACCTGGAA 120  
QY      121 CAAAGCAAGTGGCTAAATGCGCGCTTAAAGCTCTTCCCGCGCTGCTGCTGCTGCGGG 180  
Db      121 CAAAGCAAGTGGCTAAATGCGCGCTTAAAGCTCTTCCCGCGCTGCTGCTGCTGCGGG 180  
QY      181 GCAGCATCTACAGCTGCTATATAGGAGTCTTATCTCTGCTCTCTCTCTCTCTCTCTCT 240  
Db      181 GCAGCATCTACAGCTGCTATATAGGAGTCTTATCTCTGCTCTCTCTCTCTCTCTCTCT 240  
QY      241 TCCGCTTTATTTATAGGCTGGCCCTACAGGAACAACAGCTGATGTTTGAAGAACTGA 300  
Db      241 TCCGCTTTATTTATAGGCTGGCCCTACAGGAACAACAGCTGATGTTTGAAGAACTGA 300  
QY      301 CTCGTATTTGGAGACGCTACATCCAGCTCATCCCATTTCTTCTGCTGGGCTTCTACG 360  
Db      301 CTCGTATTTGGAGACGCTACATCCAGCTCATCCCATTTCTTCTGCTGGGCTTCTACG 360  
QY      361 TGAAGCTGGTGTGACCCGCTGTGTGAGAACAGTACAGAACCTGCGCGTGGCCGACCGC 420  
Db      361 TGAAGCTGGTGTGACCCGCTGTGTGAGAACAGTACAGAACCTGCGCGTGGCCGACCGC 420  
QY      421 TCATGAGCTGGTGTGCGGGCTTCTGTAAGGCAAGAGCAGCAAGCGCGCTGCGGGG 480  
Db      421 TCATGAGCTGGTGTGCGGGCTTCTGTAAGGCAAGAGCAGCAAGCGCGCTGCGGGG 480  
QY      481 GCAGGCTCATCCGCTAGCGCAACCTGGGCAACGTCATCTCGCGAGCTCAGCAACCG 540  
Db      481 GCAGGCTCATCCGCTAGCGCAACCTGGGCAACGTCATCTCGCGAGCTCAGCAACCG 540  
QY      541 CAGTCTCAAGCGCTTCCCAAGCGCCAGACGACCTGGTGCAGAGCGCTTATGACTCCG 600  
Db      541 CAGTCTCAAGCGCTTCCCAAGCGCCAGACGACCTGGTGCAGAGCGCTTATGACTCCG 600  
QY      601 CAGAACCAAGAGCTGGAGAACTGAGCCTACCAACAACATTTCTGGGTGCGCTGGG 660  
Db      601 CAGAACCAAGAGCTGGAGAACTGAGCCTACCAACAACATTTCTGGGTGCGCTGGG 660  
QY      661 TGTGTTTGGCAACCTGTCAATGAAGCGTGGCTTGAAGTCAATCCGGGACCTATCC 720  
Db      661 TGTGTTTGGCAACCTGTCAATGAAGCGTGGCTTGAAGTCAATCCGGGACCTATCC 720  
QY      721 TGTCTCAGAGCTGTCTGAGAGATGAAACACTTGGTACTAGTGTGGACACTGTATG 780  
Db      721 TGTCTCAGAGCTGTCTGAGAGATGAAACACTTGGTACTAGTGTGGACACTGTATG 780
```

```
QY      781 CCTACGACTGATTAGTATCCCACTGGGTATACACAGGTGTGACTGTGCGGGTAC 840  
Db      781 CCTACGACTGATTAGTATCCCACTGGGTATACACAGGTGTGACTGTGCGGGTAC 840  
QY      841 GCTTCTTCTGACTTGTCTAGTGTGGGGGCGAGTTTGAACCCAGCCAGGCTTACCGT 900  
Db      841 GCTTCTTCTGACTTGTCTAGTGTGGGGGCGAGTTTGAACCCAGCCAGGCTTACCGT 900  
QY      901 GCAATGAGCTGGAGACTGTGTGCGCGCTTTCACGTTCTGCACTTCTTCTATGTTG 960  
Db      901 GCAATGAGCTGGAGACTGTGTGCGCGCTTTCACGTTCTGCACTTCTTCTATGTTG 960  
QY      961 GCTGGCGGAA----- 970  
Db      961 GCTGGCGGAAAGTGGGCGCTTCCAGGGCCCTGCTGGGCTGGAGGATGGCCAGAGGGGTC 1020  
QY      971 ----- 970  
Db      1021 ATGGCCAGACGCTGTTGAGACGAGATGCAGTGTGCAAGAAAGAGTCTACAGGTAG 1080  
QY      971 ----- 970  
Db      1081 AAAGCAGCAGGCGGTGGTGGCAGACACTGTATCCAGCTACTCGGAGGCTGAGGCGAG 1140  
QY      971 ----- 970  
Db      1141 GAGAAATCGGTTGAACCGGGAGGCGGAGGTGTGGTGGCAGAGACAGCTCATCAACCCCTT 1200  
QY      998 TGGAGAGGATGATGATTTTGAACCAACTGATTTGTGACAGAAATTTGCAAGTGTG 1057  
Db      1201 TGGAGAGGATGATGATTTTGAACCAACTGATTTGTGACAGAAATTTGCAAGTGTG 1260  
QY      1058 CCTGTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1117  
Db      1261 CCTGTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320  
QY      1118 GAATTAAGCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1177  
Db      1321 GAATTAAGCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1380  
QY      1178 CTTATGAGGCTCCACCTTCAATCATGCTGAAACAAAGAGATGATGATGATGATGATG 1237  
Db      1381 CTTATGAGGCTCCACCTTCAATCATGCTGAAACAAAGAGATGATGATGATGATGATG 1440  
QY      1238 TCAGAGAGCAGAGAGATGCTCAGCTGGGCACTAATGGCGCTCTCTAGGCGCTGAGTC 1297  
Db      1441 TCAGAGAGCAGAGAGATGCTCAGCTGGGCACTAATGGCGCTCTCTAGGCGCTGAGTC 1500  
QY      1298 CCATGATACCAATCTCCAGGGCAAACTCAAGAGCAAACTACTGTGGCCCAAGAGGA 1357  
Db      1501 CCATGATACCAATCTCCAGGGCAAACTCAAGAGCAAACTACTGTGGCCCAAGAGGA 1560  
QY      1358 ATCCCTTCTCAGAGAGGCGCTGCCCAAAACCAAGCAGCAGCAACAGAGCTTAAGGG 1417  
Db      1561 ATCCCTTCTCAGAGAGGCGCTGCCCAAAACCAAGCAGCAGCAACAGAGCTTAAGGG 1620  
QY      1418 CCAGAGAGCAAAAGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 1477  
Db      1621 CCAGAGAGCAAAAGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 1680  
QY      1478 GTATCAGAGGCGAGCTACTAGTGGCCCAAGAGCGCCCTCAGGCCCACTCCATGTT 1537  
Db      1681 GTATCAGAGGCGAGCTACTAGTGGCCCAAGAGCGCCCTCAGGCCCACTCCATGTT 1740  
QY      1538 CTTCCCTTGAAGCAATCAGAGCGCTGCAAAAGCTTCAAGTGTCAAGAGCTATGACACCA 1597  
Db      1741 CTTCCCTTGAAGCAATCAGAGCGCTGCAAAAGCTTCAAGTGTCAAGAGCTATGACACCA 1800  
QY      1598 AGACAAAGCTTAAAGACTGTGATTTGGGGCCCAAAAGTTTGAATGCTGTGAGA 1657  
Db      1801 AGACAAAGCTTAAAGACTGTGATTTGGGGCCCAAAAGTTTGAATGCTGTGAGA 1860
```


[illegible]

```

repeat_region      23789..23956
                    /rpt_family="SVA"
repeat_region      26429..26707
                    /rpt_family="Alu"
repeat_region      26696
                    /rpt_family="SVA"
repeat_region      27103..27349
                    /rpt_family="Alu"
repeat_region      27323
                    /rpt_family="SVA"
repeat_region      29307
                    /rpt_family="Alu"
repeat_region      30043
                    /rpt_family="Alu"
repeat_region      29786..30017
                    /rpt_family="SVA"
repeat_region      30462..30516
                    /rpt_family="MIR"
repeat_region      31409
                    /rpt_family="Alu"
repeat_region      33005..33260
                    /rpt_family="Alu"
repeat_region      33247..33597
                    /rpt_family="L1WC2"
repeat_region      34431
                    /rpt_family="Alu"
repeat_region      34076..34411
                    /rpt_family="SVA"
repeat_region      34890
                    /rpt_family="Alu"
repeat_region      34972..35401
                    /rpt_family="Alu"
repeat_region      35392
                    /rpt_family="SVA"
repeat_region      36238..36487
                    /rpt_family="Alu"
repeat_region      36457
                    /rpt_family="SVA"
repeat_region      38855
                    /rpt_family="L1"
repeat_region      41224
                    /rpt_family="Alu"
repeat_region      41056..41213
                    /rpt_family="SVA"
repeat_region      42745..43020
                    /rpt_family="Alu"
repeat_region      42966
                    /rpt_family="SVA"
repeat_region      43312..43457
                    /rpt_family="Alu"
repeat_region      44217
                    /rpt_family="L1"
repeat_region      45533
                    /rpt_family="SVA"
repeat_region      45401
                    /rpt_family="L1R5"
repeat_region      45533
                    /rpt_family="L1R5"
repeat_region      45818
                    /rpt_family="Alu"
repeat_region      45598..45770
                    /rpt_family="SVA"
repeat_region      46004
                    /rpt_family="L1"
repeat_region      46282
                    /rpt_family="Alu"
repeat_region      46260
                    /rpt_family="SVA"
repeat_region      47583
                    /rpt_family="Alu"
repeat_region      46885..47215
                    /rpt_family="SVA"
repeat_region      47363..47371

```

```

repeat_region      /rpt_family="SVA"
                    complement(47665..47773)
repeat_region      /rpt_family="Alu"
                    complement(47792..48077)
repeat_region      /rpt_family="Alu"
                    complement(48110..48388)
repeat_region      /rpt_family="L1"
                    complement(48827..49102)
repeat_region      /rpt_family="Alu"
                    48883..49110
repeat_region      /rpt_family="SVA"
                    complement(49395..49626)
repeat_region      /rpt_family="Alu"
                    50184..50437
repeat_region      /rpt_family="Alu"

```

Query Match 28.8%; Score 642.2; DB 9; Length 142092;
 Best Local Similarity 99.5%; Pred. No. 1e-139;
 Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1202 CAGCCTGAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGAGAGAGAGATGCTCA 1261
    |||||||
DB 25481 CAGCCTGAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGAGAGAGAGATGCTCA 25422
    |||||||
QY 1262 CGCTGGCATCATTTGGCCGCTTCTAGGCGCTGCAGTCCATGATCACCATCTCCAGGCG 1321
    |||||||
DB 25421 CGCTGGCATCATTTGGCCGCTTCTAGGCGCTGCAGTCCATGATCACCATCTCCAGGCG 25362
    |||||||
QY 1322 AAACCTAGAGACCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCCAGAGGCGCTGCC 1381
    |||||||
DB 25361 AAACCTAGAGACCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCCAGAGGCGCTGCC 25302
    |||||||
QY 1382 CAAAACCCAAAGCGAGCCCAACAGAAAGCTTAGGGCCAGAGACAAAGGCGCTGGAA 1441
    |||||||
DB 25301 CAAAACCCAAAGCGAGCCCAACAGAAAGCTTAGGGCCAGAGACAAAGGCGCTGGAA 25242
    |||||||
QY 1442 GCTTAAGGCTGTGGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCGAGCTACTACAG 1501
    |||||||
DB 25241 GCTTAAGGCTGTGGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCGAGCTACTACAG 25182
    |||||||
QY 1502 TGCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTTAGAACCATCAGGCC 1561
    |||||||
DB 25181 TGCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTTAGAACCATCAGGCC 25122
    |||||||
QY 1562 GTCAAGCTTTCACAGTGTCCAGAGCAAGAGACCAAGAAAGCTTAAGACTGTGAG 1621
    |||||||
DB 25121 GTCAAGCTTTCACAGTGTCCAGAGCAAGAGACCAAGAAAGCTTAAGACTGTGAG 25062
    |||||||
QY 1622 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGCGCTTGATGAGCA 1681
    |||||||
DB 25061 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGCGCTTGATGAGCA 25002
    |||||||
QY 1682 CCCAGAAGTATCTCAAGTGAAGAGAAACTGTGGAGTTTAACCTAGCGATATGCCAGA 1741
    |||||||
DB 25001 CCCAGAAGTATCTCAAGTGAAGAGAAACTGTGGAGTTTAACCTAGCGATATGCCAGA 24942
    |||||||
QY 1742 GATCCCCGAAATCACCCTCAAGACCTTGAACAATCAGCAACCAACATACACTAC 1801
    |||||||
DB 24941 GATCCCCGAAATCACCCTCAAGACCTTGAACAATCAGCAACCAACATACACTAC 24882
    |||||||
QY 1802 ACTCAAGATCAGATGATCTTATTTGGGCGCTTGGAAGACGAGGATG 1848
    |||||||
DB 24881 ACTCAAGATCAGATGATCTTATTTGGGCGCTTGGAAGACGAGGATG 24835
    |||||||

```

RESULT 6
 AC087451/ 163915 bp DNA 1linear HTG 04-JUL-2001
 LOCUS Homo sapiens chromosome 11 clone Rb11-810p12 map 11, WORKING DRAFT
 DEFINITION SEQUENCE 7 unordered pieces.
 ACCESSION AC087451
 VERSION AC087451.2 GI:14595836
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

| | | | |
|---------------------------|--------|---|------------------------|
| QY | 1742 | GATCCCGGAAATACCTCAAGAAAGAAAGAACTTTGAGCAATACCAACCAACCAATACACACTAC | 1801 |
| Db | 150890 | GATCCCGGAAATACCTCAAGAAAGAAAGAACTTTGAGCAATACCAACCAACCAATACACACTAC | 1508 |
| QY | 1802 | ACTCAAGATCACATGATGATCTTATTTGGGCGCTTGAAACAGGATG | 1848 |
| Db | 150830 | ACTCAAGATCACATGATGATCTTATTTGGGCGCTTGAAACAGGATG | 150784 |
| RESULT 7 | | | |
| AP003733 | | 166867 bp | DNA |
| LOCUS | | | linear PRI 08-DEC-2001 |
| DEFINITION | | Homo sapiens genomic DNA, chromosome 11q, clone:RP11-810P12, | |
| ACCESSION | | Complete sequence. | |
| VERSION | | AP003733 | |
| KEYWORDS | | AP003733.4 GI:17426128 | |
| SOURCE | | HTG. | |
| ORGANISM | | Homo sapiens | |
| REFERENCE | | Homo sapiens DNA, clone:RP11-810P12. | |
| AUTHORS | | Eukaryote, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | |
| TITLE | | 1 | |
| JOURNAL | | Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y. | |
| REFERENCE | | Homo sapiens genomic DNA | |
| AUTHORS | | Published Only in Database (2001) | |
| TITLE | | 2 (bases 1 to 166867) | |
| JOURNAL | | Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y. | |
| REFERENCE | | Direct Submission | |
| AUTHORS | | Submitted (11-JUN-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Tsukuba-shi, Ibaraki, 305-8565, Japan (E-mail:hattori@gs.c.riken.go.jp, URL:http://hnp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170) | |
| COMMENT | | On Dec 7, 2001 this sequence version replaced gi:17026124. | |
| FEATURES | | location/Qualifiers | |
| SOURCE | | 1. 166867 | |
| | | /organism="Homo sapiens" | |
| | | /db_xref="taxon:9606" | |
| | | /chromosome="11" | |
| | | /map="11q" | |
| | | /clone="RP11-810P12" | |
| BASE COUNT | | 44705 a 38465 c 39268 g 44429 t | |
| ORIGIN | | | |
| Query Match | | 28.8%; Score 642.2; DB 9; Length 166867; | |
| Best Local Similarity | | 99.5%; Pred. No. 1e-139; | |
| Matches 644; Conservative | | 0; Mismatches 3; Indels 0; Gaps 0; | |
| QY | 1202 | CAGCTGAAACAAAGAGAGATGAGATTCACAGCCCAATAGAGAGACGAGAGATGCTCA | 1261 |
| Db | 12486 | CAGCTGAAACAAAGAGAGATGAGATTCACAGCCCAATAGAGAGACGAGAGATGCTCA | 1254 |
| QY | 1262 | CGCTGGCATATTTGGCGCGCTTCAGAGGCTGCAAGTCCCATGATCCATCTCCAGGCG | 1321 |
| Db | 12546 | CGCTGGCATATTTGGCGCGCTTCAGAGGCTGCAAGTCCCATGATCCATCTCCAGGCG | 1260 |
| QY | 1332 | AAACTCAAGACCAAACTACTGTGTGGCCCAAGAGGGAATCCCTTCACAGAGGCGCTGC | 1381 |
| Db | 12606 | AAACTCAAGACCAAACTACTGTGTGGCCCAAGAGGGAATCCCTTCACAGAGGCGCTGC | 1266 |
| QY | 1382 | CAAAAACCAAGAGGACGCAACAAAGAAAGAGTAAAGGGCCAGAAAGACAAAGGCGCTGAA | 1441 |
| Db | 12666 | CAAAAACCAAGAGGACGCAACAAAGAAAGAGTAAAGGGCCAGAAAGACAAAGGCGCTGAA | 1272 |
| QY | 1442 | GCTTAAGGCTGTGAGAGGCTTCAAGTCTGGGCCAATGTATGAGAGGCGAGGCTACTACAG | 1501 |
| Db | 12726 | GCTTAAGGCTGTGAGAGGCTTCAAGTCTGGGCCAATGTATGAGAGGCGAGGCTACTACAG | 1278 |
| QY | 1502 | TGCCCCACAGACGCCCTCAGGCCCACTCCCATGTTCTTCCCTTAGAACATCAGCGCC | 1561 |

| Db | 12786 | TGCCCCACAGACGCCCCCTCAGGCCACCCACATGTTCTTCGCCCTAGAACATCAAGCGCC | 12845 |
|------------|--|--|--------|
| QY | 1562 | GTCAAAGCTTCACAGTGTCCACAGCGCATGACACCAACAAAGCTTTAAAGACTGTGAG | 1621 |
| Db | 12846 | GTCAAAGCTTCACAGTGTCCACAGCGCATGACACCAACAAAGCTTTAAAGACTGTGAG | 12905 |
| QY | 1622 | TTCTGGGGGGCCCAAGAAAAGTTTGAATTTGGCTCTCAGAGAGCGATGGGGCCCTGTGATGGAGCA | 1681 |
| Db | 12906 | TTCTGGGGGGCCCAAGAAAAGTTTGAATTTGGCTCTCAGAGAGCGATGGGGCCCTGTGATGGAGCA | 12965 |
| QY | 1682 | CCCAAGAGTATCTCAAGTGTGAGAGAGAAAACGTGTGAGATTAACTGACGAGATATGCGCAGA | 1741 |
| Db | 12966 | CCCAAGAGTATCTCAAGTGTGAGAGAGAAAACGTGTGAGATTAACTGACGAGATATGCGCAGA | 13025 |
| QY | 1742 | GATCCCGGAAAATCACCCTCAAAAACCTTTGGAACCAATACCAACCAATACACACTAC | 1801 |
| Db | 13026 | GATCCCGGAAAATCACCCTCAAAAACCTTTGGAACCAATACCAACCAATACACACTAC | 13085 |
| QY | 1802 | ACTCAAGATCACATGGATCCTTATTGGGCGCTTGAAAACAGGATG | 1848 |
| Db | 13086 | ACTCAAGATCACATGGATCCTTATTGGGCGCTTGAAAACAGGATG | 13132 |
| RESULT 8 | | | |
| AC004228 | 196080 bp | DNA | linear |
| LOCUS | Homo sapiens Chromosome 11q12.2 PAC clone pD519013 containing | | |
| DEFINITION | human gene for ferritin heavy chain (FTH), complete sequence. | | |
| AC004228 | | | |
| VERSION | AC004228.2 | GI:4263838 | |
| KEYWORDS | HTG. | | |
| SOURCE | Homo sapiens. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | |
| AUTHORS | 1 (bases 1 to 196080) | | |
| | Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brlgnac,S., Bumesler,T., Davis,C., English,C., Franklin,T.L., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osbourne-Lawrence,S., Schageman,J., Schultz,R.A., Stinson,S., Syed,M., Valenzuela,D., Ward,T. and Wilson,R. | | |
| TITLE | HRGS Submision | | |
| JOURNAL | Unpublished | | |
| REFERENCE | 2 (bases 1 to 196080) | | |
| AUTHORS | Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brlgnac,S., Bumesler,T., Davis,C., English,C., Franklin,T.L., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Hinson,S., Naryanaswamy,U., Newton,J., O'Brien,K., Patel,P., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (26-FEB-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA | | |
| REFERENCE | 3 (bases 1 to 196080) | | |
| AUTHORS | Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buelther,J., Butler,C., Card,P., desaillobert,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osbourne-Lawrence,S., Schageman,J., Schultz,R.A., Stinson,S., Waller,K. and Ward,T. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (24-FEB-1999) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA | | |
| COMMENT | On Feb 24, 1999 this sequence version replaced g1:2911733. IMPORTANT: This submission contains the entire insert of clone pD519013. pD519013 comes from the RPI-3 PAC library constructed at the Roswell Park Cancer Institute by the Pacer de Jong group. CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11p12.2 Best's disease region mapped between STS D11S461 and EST ANNAK. This region spans over 1.5 Mbp. | | |

MARKER CONFIRMATION: EST: FTH (ferritin heavy chain mRNA), STSS
D11659 and WI-7524
MAPED CLONE OVERLAP: PACs PDU66a11 and PDU756b9.

FEATURES
Location/Qualifiers

| source | 1.196060 | /organism="Homo sapiens" | /db_xref="taxon:9606" | complement(2210..2482) |
|---------------|--------------------------|--------------------------|-----------------------|------------------------|
| repeat_region | 3457..3743 | /rpt_family="Alu" | | |
| repeat_region | /rpt_family="Alu" | | | |
| repeat_region | complement(20145..20758) | /rpt_family="Alu" | | |
| repeat_region | complement(23194..23599) | /rpt_family="Alu" | | |
| repeat_region | complement(27367..27685) | /rpt_family="Alu" | | |
| repeat_region | 28047..28457 | /rpt_family="Alu" | | |
| repeat_region | complement(29390..29565) | /rpt_family="Alu" | | |
| repeat_region | complement(31076..31255) | /rpt_family="Alu" | | |
| repeat_region | complement(34948..35045) | /rpt_family="Alu" | | |
| repeat_region | 35232..35538 | /rpt_family="Alu" | | |
| repeat_region | complement(36242..36540) | /rpt_family="Alu" | | |
| repeat_region | 37238..37522 | /rpt_family="Alu" | | |
| repeat_region | complement(38555..38650) | /rpt_family="Alu" | | |
| repeat_region | complement(40483..41164) | /rpt_family="Alu" | | |
| repeat_region | 42042..42116 | /rpt_family="MIR" | | |
| repeat_region | 42677..42817 | /rpt_family="L1" | | |
| repeat_region | 42856..43157 | /rpt_family="L1" | | |
| repeat_region | 43148..43380 | /rpt_family="L1" | | |
| repeat_region | complement(43734..43999) | /rpt_family="L1" | | |
| repeat_region | complement(46998..47169) | /rpt_family="MIR" | | |
| repeat_region | complement(48048..48346) | /rpt_family="Alu" | | |
| repeat_region | 48482..48737 | /rpt_family="Alu" | | |
| repeat_region | complement(48893..49037) | /rpt_family="L1" | | |
| repeat_region | 49118..49292 | /rpt_family="Alu" | | |
| repeat_region | 49791..50059 | /rpt_family="Alu" | | |
| repeat_region | 50436..50501 | /rpt_family="Alu" | | |
| repeat_region | complement(51147..51441) | /rpt_family="Alu" | | |
| repeat_region | complement(52700..52979) | /rpt_family="Alu" | | |
| repeat_region | 53281..53568 | /rpt_family="Alu" | | |
| repeat_region | 54071..54204 | /rpt_family="Alu" | | |
| repeat_region | 54506..54793 | /rpt_family="Alu" | | |
| repeat_region | 54807..55061 | /rpt_family="Alu" | | |
| repeat_region | complement(64533..64826) | /rpt_family="Alu" | | |

| | | | | |
|---------------|----------------------------|---------------------|--|--|
| repeat_region | complement(65245..65783) | /rpt_family="Alu" | | |
| repeat_region | complement(68465..68732) | /rpt_family="Alu" | | |
| repeat_region | complement(70191..70476) | /rpt_family="Alu" | | |
| repeat_region | 70560..70692 | /rpt_family="Alu" | | |
| repeat_region | 72294..72470 | /rpt_family="Alu" | | |
| repeat_region | complement(72800..73095) | /rpt_family="Alu" | | |
| repeat_region | 73424..73695 | /rpt_family="Alu" | | |
| repeat_region | 73784..74071 | /rpt_family="Alu" | | |
| repeat_region | 74652..74935 | /rpt_family="Alu" | | |
| repeat_region | complement(75265..75514) | /rpt_family="Alu" | | |
| repeat_region | complement(76086..76353) | /rpt_family="Alu" | | |
| repeat_region | complement(78038..78312) | /rpt_family="Alu" | | |
| repeat_region | 78513..78734 | /rpt_family="Alu" | | |
| repeat_region | 78866..79180 | /rpt_family="Alu" | | |
| repeat_region | complement(79384..79673) | /rpt_family="Alu" | | |
| repeat_region | complement(81398..81690) | /rpt_family="Alu" | | |
| repeat_region | 82013..82288 | /rpt_family="Alu" | | |
| repeat_region | complement(83419..83600) | /rpt_family="MIR" | | |
| repeat_region | complement(87320..87624) | /rpt_family="Alu" | | |
| repeat_region | complement(87710..87987) | /rpt_family="Alu" | | |
| repeat_region | complement(88186..88472) | /rpt_family="Alu" | | |
| repeat_region | 88822..89049 | /rpt_family="Alu" | | |
| repeat_region | 89224..89308 | /rpt_family="MER20" | | |
| repeat_region | 89729..90487 | /rpt_family="Alu" | | |
| repeat_region | 90504..90618 | /rpt_family="Alu" | | |
| repeat_region | complement(93313..93952) | /rpt_family="Alu" | | |
| repeat_region | 96938..97517 | /rpt_family="Alu" | | |
| repeat_region | complement(104199..104514) | /rpt_family="Alu" | | |
| repeat_region | complement(109738..109979) | /rpt_family="Alu" | | |
| repeat_region | 110022..110141 | /rpt_family="Alu" | | |
| repeat_region | complement(111091..112016) | /rpt_family="Alu" | | |
| repeat_region | 113210..113419 | /rpt_family="L1" | | |
| repeat_region | complement(114951..115005) | /rpt_family="MIR" | | |
| repeat_region | 115685..115791 | /rpt_family="MIR" | | |
| repeat_region | 116870..117005 | /rpt_family="MIR" | | |
| repeat_region | complement(117877..118172) | /rpt_family="Alu" | | |
| repeat_region | 118352..118600 | /rpt_family="Alu" | | |

repeat_region /rpl_family="Alu"
118679..118813
repeat_region /rpl_family="Alu"
complement(121233..121520)
repeat_region /rpl_family="Alu"
122658..123071

Query Match 28.8%; Score 642.2; DB 9; Length 196080;
Best Local Similarity 99.5%; Pred. No. 1.1e-139;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1202 CAGCTGAAACAAGAGAGATGAGATTCAGCCCAATCAGAGAGAGAGAGATGCTCA 1261
DB 191325 CAGCTGAAACAAGAGAGATGAGATTCAGCCCAATCAGAGAGAGAGAGATGCTCA 191384
QY 1262 CGCTGGCATCATTTGGCCCTCTTCTAGGCTGAGTCCATGATACATCCCTCCAGAGGC 1321
DB 191385 CGCTGGCATCATTTGGCCCTCTTCTAGGCTGAGTCCATGATACATCCCTCCAGAGGC 191444
QY 1322 AAATCAAGACCAAACTACTGTGGCCCAAGAGAGAGATCCCTTCCAGAGAGGCCTGCC 1381
DB 191445 AAATCAAGACCAAACTACTGTGGCCCAAGAGAGAGATCCCTTCCAGAGAGGCCTGCC 191504
QY 1382 CAAAACCAAG 1441
DB 191505 CAAAACCAAG 191564
QY 1442 GCTTAAGGCTGTGAG 1501
DB 191565 GCTTAAGGCTGTGAG 191624
QY 1502 TGCCCAAG 1561
DB 191625 TGCCCAAG 191684
QY 1562 GTCAAGGCTGTGAG 1621
DB 191685 GTCAAGGCTGTGAG 191744
QY 1622 TTCTGGGGCCCAAG 1681
DB 191745 TTCTGGGGCCCAAG 191804
QY 1682 CCCAAGATATCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1741
DB 191805 CCCAAGATATCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 191864
QY 1742 GATCCCCGAAATATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1801
DB 191865 GATCCCCGAAATATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 191924
QY 1802 ACTCAAGATATCACTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1848
DB 191925 ACTCAAGATATCACTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 191971

RESULT 9
HSYMD2P10 706 bp DNA linear PRI 22-OCT-1998
DEFINITION Homo sapiens vitelliform macular dystrophy protein (VMD2) gene,
exon 10.
ACCESSION AF073499
VERSION AF073499.1 GI:3598873
KEYWORDS
SEGMENT 10 of 11
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 706)
AUTHORS Stohr,H., Marguardt,A., Rivera,A., Cooper,P.R., Nowak,N.J.,
Shows,T.B., Gerhard,D.S. and Weber,B.H.
A gene map of the Best's vitelliform macular dystrophy region in

JOURNAL Chromosome 11q12-q13.1
MEDLINE Genome Res. 8 (1), 48-56 (1998)
PUBMED 98112782
REFERENCE 9445487
AUTHORS 2 (bases 1 to 706)
Marguardt,A., Stohr,H., Passmore,L.A., Kramer,F., Rivera,A. and
Weber,B.H.

TITLE Mutations in a novel gene, VMD2, encoding a protein of unknown
properties cause juvenile-onset vitelliform macular dystrophy
(Best's disease).
Hum. Mol. Genet. 7 (9), 1517-1525 (1998)

JOURNAL 3 (bases 1 to 706)
MEDLINE Marguardt,A., Stohr,H., Passmore,L., Kraemer,F., Rivera,A. and
PUBMED 98367043
REFERENCE 9700209
AUTHORS Weber,B.H.F.

TITLE Direct Submission
JOURNAL Submitted (23-JUN-1998) Human Genetics, University, Biozentrum, Am
Hubland, Wuerzburg 97074, Germany
FEATURES Location/Qualifiers

source 1..706
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q12-q13.1"
/clone="PAC d446a11"
36..674
/gene="VMD2"
/number=10

BASE COUNT 204 a 209 c 164 g 129 t

ORIGIN

Query Match 28.7%; Score 640.6; DB 9; Length 706;
Best Local Similarity 99.4%; Pred. No. 1.1e-139;
Matches 643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1202 CAGCTGAAACAAGAGAGATGAGATTCAGCCCAATCAGAGAGAGAGAGATGCTCA 1261
DB 33 CAGCTGAAACAAGAGAGATGAGATTCAGCCCAATCAGAGAGAGAGAGATGCTCA 92
QY 1262 CGCTGGCATCATTTGGCCCTCTTCTAGGCTGAGTCCATGATACATCCCTCCAGAGGC 1321
DB 93 CGCTGGCATCATTTGGCCCTCTTCTAGGCTGAGTCCATGATACATCCCTCCAGAGGC 152
QY 1322 AAATCAAGACCAAACTACTGTGGCCCAAGAGAGAGATCCCTTCCAGAGAGGCCTGCC 1381
DB 133 AAATCAAGACCAAACTACTGTGGCCCAAGAGAGAGATCCCTTCCAGAGAGGCCTGCC 212
QY 1382 CAAAACCAAG 1441
DB 213 CAAAACCAAG 272
QY 1442 GCTTAAGGCTGTGAG 1501
DB 273 GCTTAAGGCTGTGAG 332
QY 1502 TGCCCAAG 1561
DB 333 TGCCCAAG 392
QY 1562 GTCAAGGCTGTGAG 1621
DB 393 GTCAAGGCTGTGAG 452
QY 1622 TTCTGGGGCCCAAG 1681
DB 453 TTCTGGGGCCCAAG 512
QY 1682 CCCAAGATATCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1741
DB 513 CCCAAGATATCTCAAGTGAAG 572
QY 1742 GATCCCCGAAATATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1801

Db 573 GATCCCCGAAATACCTCTCAAGAACCTTTGGACATACCAACCAACATACACACTAC 632

Db 1802 ACTCAAGATCAGATGATCCCTTATTGGGCGCTTGGAAAAAGAGATG 1848

Db 633 ACTCAAGATCAGATGATCCCTTATTGGGCGCTTGGAAAAAGAGCTTG 679

| RESULT 10 | | | | | |
|-----------|------------|-----------------------------------|------------|--------|--------------------|
| AC003025 | LOCUS | | | | |
| AC003025 | DEFINITION | 112309 bp | DNA | linear | PRI 23-JUL-1996 |
| | ACCESSION | Human Chromosome 1p12.2 PAC clone | pdf66a11 | | complete sequence. |
| | VERSION | AC003025 | | | |
| | KEYWORDS | AC003025.1 | GI:3337308 | | |
| | SOURCE | HTG. | | | |
| | ORGANISM | Homo sapiens. | | | |
| | | Homo sapiens | | | |

REFERENCE
AUTHORS
1 (bases 1 to 112309)
Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Basil, M.

| TITLE | JOURNAL | REFERENCE | AUTHORS |
|-----------------|-------------|-----------------------|---|
| HTGS Submission | Unpublished | 2 (bases 1 to 112309) | Evans, G. A., Athanassiou, M., Bradbury, P., Bryant, S. |

| REFERENCE | AUTHORS | TITLE | JOURNAL |
|-----------------------|--|-------------------|--|
| 2 (bases 1 to 112309) | Evans, G.A., Athanasiou, M., Bradbury, P., Briñac, S., Bumeister, R., Davis, C., English, C., Franklin, T.L., Garner, H.R., Gordon, M., Gowan, G., Grant, O., Hanner, L., Harris, J., Hinson, S., Naryanaswamy, U., Newton, J., O'Brien, K., Oliver, T., Patel, P., Probst, S., Rayner, S., Schageman, J., Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T. and Wilson, R. | Direct Submission | Submitted (21-Oct-1997) Genome Science and Technology Center, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA |
| 3 (bases 1 to 112309) | Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Bacht, M. | | |

English, C., Ethridge, S., Garner, H.R., Gee, V., Gordon, M., Gotway, G., Grant, O., Hamner, L., Joslin, J., Lewis, E., Loo, H., Loo, K.N., Major, T., McFarland, J., Newton, J., Osborne-Lawrence, E., Schagman, J., Schultz, R.A., Stimson, S., Syed, M. and Ward, T.

COMMENT

ON OUR 23, 1998 this sequence version replaced 91.2554967. IMPORTANT: This submission contains the entire insert of clone PDJ466all. PDJ466all comes from a PAC library constructed at the Roswell Park Cancer Institute by the Pletier de Jong group. This clone has been finished according to strict quality criteria and attempts have been made to resolve all base calling problems such as compressions and repetitive elements. The expected Phred/Phrap calculated errors/100k is 0.18. In addition, this sequence has been finished such that 99.9% of consensus base calls consist of either double-stranded coverage or 2 types of labeling chemistry on one strand.

Further information regarding the map of this region or annotation of pDj46a1 can be found at <http://gestic.swmed.edu/chromosomel.htm>. CHROMOSOMAL LOCUS: This PAC clone comes from 11P12.2 Best's disease region mapped between STS D11S461 and EST HANK. This region spans over 1.5 Mbp. CLONE CONFIRMATION: EST, ETH (retinitin heavy chain mRNA) MAPPED CLONE OVERLAP: HTGS-submitted PAC clones pdu519c13 and pdu756b9.

```

FEATURES
source      Location/Qualifiers
1. 112309   /organism="Homo sapiens"

```

| | |
|---------------|---|
| repeat_region | /db_xref="taxon:9606" complement(1341..4656) |
| repeat_region | /rpt_family="Alu" complement(8880..10120) |
| repeat_region | /rpt_family="Alu" 10163..10282 |
| repeat_region | /rpt_family="Alu" complement(11232..13351..13560) |
| repeat_region | /rpt_family="Alu" complement(15092..15146) |
| repeat_region | /rpt_family="MIR" 15826..15932 |
| repeat_region | /rpt_family="MIR" 17011..17146 |
| repeat_region | /rpt_family="MIR" complement(18018..18493..18741) |
| repeat_region | /rpt_family="Alu" 18820..18954 |
| repeat_region | /rpt_family="Alu" complement(21374..22809..23212) |
| repeat_region | /rpt_family="Alu" 23499..23752 |
| repeat_region | /rpt_family="Alu" 24118..24472 |
| repeat_region | /rpt_family="Alu" 25005..25113 |
| repeat_region | /rpt_family="MIR" 31160..31442 |
| repeat_region | /rpt_family="Alu" complement(33251..35770..36025) |
| repeat_region | /rpt_family="MIR" 38958..39145 |
| repeat_region | /rpt_family="Alu" 44878..44880 |
| repeat_region | /rpt_family="Alu" 49769..50071 |
| repeat_region | /rpt_family="Alu" complement(50172..50902..50982) |
| repeat_region | /rpt_family="MIR" 51038..51659 |
| repeat_region | /rpt_family="Alu" 54494..54769 |
| repeat_region | /rpt_family="Alu" 55529..55623 |
| repeat_region | /rpt_family="TBE1" complement(55691..55967) |
| repeat_region | /rpt_family="Alu" complement(56049..56257) |
| repeat_region | /rpt_family="Alu" 56721..57100 |
| repeat_region | /rpt_family="TBE1" 58402..58680 |
| repeat_region | /rpt_family="Alu" complement(58748..59345) |
| repeat_region | /rpt_family="Alu" 59678..59904 |
| repeat_region | /rpt_family="Alu" 61149..61424 |
| repeat_region | /rpt_family="Alu" complement(61930..62242) |
| repeat_region | /rpt_family="Alu" 62276..62710 |
| repeat_region | /rpt_family="MUT1" |

```

repeat_region      62915..63118
                    /rpt_family="Alu"
repeat_region      complement(63349)..63455
                    /rpt_family="Alu"
repeat_region      64339..64614
                    /rpt_family="Alu"
repeat_region      complement(64707)..64984
                    /rpt_family="Alu"
repeat_region      complement(65596)..65811
                    /rpt_family="Alu"
repeat_region      complement(66500)..66753
                    /rpt_family="Alu"
repeat_region      67311..67542
                    /rpt_family="Alu"
repeat_region      67835..68119
                    /rpt_family="Alu"
repeat_region      68860..69150
                    /rpt_family="Alu"
repeat_region      69164..69272
                    /rpt_family="Alu"
repeat_region      69354..70245
                    /rpt_family="Alu"
repeat_region      70655..70939
                    /rpt_family="Alu"
repeat_region      70933..71075
                    /rpt_family="L1"
repeat_region      71119..71395
                    /rpt_family="Alu"
repeat_region      71404..71530
                    /rpt_family="LTR5"
repeat_region      71536..71637
                    /rpt_family="LTR5"
repeat_region      72720..72948
                    /rpt_family="L1"
repeat_region      complement(73488)..73633
                    /rpt_family="Alu"
repeat_region      complement(73925)..74200
                    /rpt_family="Alu"
repeat_region      75721..75955
                    /rpt_family="Alu"
repeat_region      77185..77449
                    /rpt_family="Alu"
repeat_region      78093..78222
                    /rpt_family="L1"
repeat_region      78227..78294
                    /rpt_family="L1"
repeat_region      complement(80458)..80707
                    /rpt_family="Alu"
repeat_region      complement(81544)..81974
                    /rpt_family="Alu"
repeat_region      82055..82305
                    /rpt_family="Alu"
repeat_region      82514..83096
                    /rpt_family="Alu"
repeat_region      complement(83685)..83940
                    /rpt_family="Alu"
repeat_region      85536..85850
                    /rpt_family="Alu"
repeat_region      complement(86429)..86483
                    /rpt_family="MIR"
repeat_region      86902..87183
                    /rpt_family="Alu"
repeat_region      87546..87792
                    /rpt_family="Alu"

```

Query Match 28.7% Score 640.6; DB 9; Length 112309;
 Best Local Similarity 99.4%; Pred. No. 2.3e-139;
 Matches 643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1202 CAGCCTGACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGACGAGAGATGCTCA 1261
 DB 91470 CAGCCTGACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGACGAGAGATGCTCA 91529

```

QY 1262 CGCTGGATCATTTGGCGCTTCTCCTAGGCGCTGACATCCATGATCACCATCTCCAGGCG 1321
DB 91530 CGCTGGATCATTTGGCGCTTCTCCTAGGCGCTGACATCCATGATCACCATCTCCAGGCG 91589
QY 1322 AAATCAAGGACCAAACTACTGTGGGCCCAAGAGGAAATCCCTTCCACGAGGCGCTGCC 1381
DB 91590 AAATCAAGGACCAAACTACTGTGGGCCCAAGAGGAAATCCCTTCCACGAGGCGCTGCC 91649
QY 1382 CAAAACCCACAGGCGACCAAAAGAAAGCTTAGGCGCCAGAGACAAACAAAGGCGTGGAA 1441
DB 91650 CAAAACCCACAGGCGACCAAAAGAAAGCTTAGGCGCCAGAGACAAACAAAGGCGTGGAA 91709
QY 1442 GCTTAAGGCTGTGAGCGCTTCAAGTGTGGCGCCACTGTATGAGAGCGGCGCTACTAGG 1501
DB 91710 GCTTAAGGCTGTGAGCGCTTCAAGTGTGGCGCCACTGTATGAGAGCGGCGCTACTAGG 91769
QY 1502 TGCCCAACAGAGCGCCCTCAGCCCTCAGCCCTCAGCTGTCTTCCCTAGAACCATCAGCGCC 1561
DB 91770 TGCCCAACAGAGCGCCCTCAGCCCTCAGCCCTCAGCTGTCTTCCCTAGAACCATCAGCGCC 91829
QY 1562 GTCAAGCTTCAAGTGTACAGGCGATGACACCAACAAAGAAAGCTTAAGACTGTGAG 1621
DB 91830 GTCAAGCTTCAAGTGTACAGGCGATGACACCAACAAAGAAAGCTTAAGACTGTGAG 91889
QY 1622 TTCTGGGGCCAAAGAAAGTTTGAATGCTGTACAGAGCGATGAGGCGCTGATGAGCA 1681
DB 91890 TTCTGGGGCCAAAGAAAGTTTGAATGCTGTACAGAGCGATGAGGCGCTGATGAGCA 91949
QY 1682 CCCAAGATATCTCAAGTGAAGAGAGAAACGTGTGAGTTTAACCTGACGATATGCGAGA 1741
DB 91950 CCCAAGATATCTCAAGTGAAGAGAGAAACGTGTGAGTTTAACCTGACGATATGCGAGA 92009
QY 1742 GATCCCGGAAATATACCTCCTCAAGAAACCTTGTGAAACAAACACCAACATACACATAC 1801
DB 92010 GATCCCGGAAATATACCTCCTCAAGAAACCTTGTGAAACAAACACCAACATACACATAC 92069
QY 1802 ACTCAAGATCAGATGATCTTATTTGGGCGCTTGTGAAACAGAGGATG 1848
DB 92070 ACTCAAGATCAGATGATCTTATTTGGGCGCTTGTGAAACAGAGGATG 92116

```

RESULT 11
 AC084857
 LOCUS
 DEFINITION Homo sapiens chromosome 11, clone CTD-3231N5 map 11, WORKING DRAFT
 ACCESSION AC084857
 VERSION AC084857.1 GI:11276215
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abramson,H., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,
 Boukhalter,B., Brown,A., Burkett,G., Campilongo,A., Castle,A.,
 Choquel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
 Dekrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
 Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
 Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
 Iliev,I., Johnson,R., Jones,C., Kand,L., Karats,A., LaRoque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
 Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
 McNeeters,R., Meldrum,J., Menus,L., Mihova,T., Mlenga,V.,
 Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
 O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
 Piere,N., Pisanl,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
 Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,

Db 92912 ACTCAAGATCATGATGATCTTATGTGGCTTGGAAGAGGCTGTG 92958

RESULT 12
AF440756 1908 bp mRNA linear PRI 12-JUL-2002
LOCUS Homo sapiens vitelliform macular dystrophy 2-like protein 1 mRNA,
DEFINITION complete cds.
ACCESSION AF440756
VERSION AF440756.1 GI:21734839

KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Marguardt, A., Stohr, H., Passmore, L. A., Kramer, F., Rivera, A. and
Weber, B. H.
TITLE Mutations in a novel gene, VMD2, encoding a protein of unknown
properties cause juvenile-onset vitelliform macular dystrophy
(Best's disease)
JOURNAL Hum. Mol. Genet. 7 (9), 1517-1525 (1998)
MEDLINE 98367043
PUBMED 9702029

REFERENCE
AUTHORS Stohr, H., Marguardt, A., Nanda, I., Schmid, M. and Weber, B. H.
TITLE Three novel human VMD2-like genes are members of the evolutionary
highly conserved RFP-TM family
JOURNAL Eur. J. Hum. Genet. 10 (4), 281-284 (2002)
MEDLINE 12032749
PUBMED 12032738

REFERENCE
AUTHORS Stohr, H., Marguardt, A. and Weber, B. H.
TITLE Three novel human VMD2-like protein genes are members of the
evolutionary highly conserved RFP family
JOURNAL Unpublished
AUTHORS 4 (bases 1 to 1908)
Stohr, H., Marguardt, A. and Weber, B. H.
TITLE Direct Submission
JOURNAL Submitted (30-OCT-2001) Human Genetics, University of Wuerzburg,
Biozentrum, Am Hubland, Wuerzburg 97074, Germany

FEATURES
source location/Qualifiers
1. C.1908
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.12-p13.2"
1. 1530
/note="VMD2L1: putative membrane protein; Pfam01062;
worm_family_8"
/codon_start=1
/product="vitelliform macular dystrophy 2-like protein 1"
/protein_id="AA076995.1"
/db_xref="GI:21734840"
/translation="MTVTYARVANAARFGGSQLLLMWSIYKLLMRELLCGLGYM
ALSAARFVLTGEGKRYFEKLVYCDQYASLIPVSFVGLGYVLVYVNRMSOYLCHPL
PALMKVAVGTVGRDGRGLRYRLRLAYGLSAVLLRSVAVRKRPTIDHYVEA
GKTRERKKEFNLSYNKYVPCWFSMLAARREGIRNSALXILLLEIANFR
GKGMLEHNDMSVPLVYQVYVIAKSYSLACLIGROPIDRPOGKDKHDLDICVIF
TILQFFRYGWLKVAEQLINPREGDDDEFINFLIDRNFOVSLADVEMDIDLAILEK
DLYWDAEAAAPYATATVAFQLRQPSFGSTFDITLAKEDMQFRLDGLDGMKEAEGD
FLORLLPAGAGVAGGPIGRRLSFLLRKNCVSEASTKAGSCAVVPEGAAPEACSGD
PLIDPLPEPEAPPPAGPEPLTLIPGPVEFISIVTMPGPRPAPMLPSPIGEEENL
A"

CDS
BASE COUNT 337 a 604 c 551 g 416 t

ORIGIN

Query Match 25.9%; Score 576.8; DB 9; Length 1908;
Best Local Similarity 70.0%; Pred. No. 1.3e-124;
Matches 792; Conservative 0; Mismatches 337; Indels 3; Gaps 1;

QY 105 ATGACATCACTTACACAGCCAGTGGCTAATGCCCGCTTACGCTTCCTCCGCGCTG 164

Db 1 ATGACGCTACACTACACAGCCAGTGGCGAAGCCCGCTTCGGTGGCTTCCAGCTG 60

QY 165 CTGCTGTGTCGGCGGGCAGCATCTACAGCTGCTATATGGAGTTCATCTCTG 224

Db 61 CTGCTACTGTGGCTGGAGCATCTACAACTCTGTGGAGAGCTCTGCTCTCTT 120

QY 225 CTCGCTACATCATCATCCGTTTATTTATAGCTGGCCCTCAGAGGAACAACAGCTG 284

Db 121 GGGTTCACATGCGGCTGAGTGTGCTTACCGCTTGTGTGACCGAAGGCAAGCGC 180

QY 285 ATGTTGAGAAACAGACTGTGTATGGCAGCTACATCCAGCTCATCCCATTTCTTC 344

Db 181 TACTTCAGAAAGCTGTGTGATTTATGTGACCAAGTATGCCAGCTCATCCCTCTCTC 240

QY 345 GTGCTGGCTTCTACGTGACGCTGTGTGACCCGCTGTGTGGAACAGTACAGAACTG 404

Db 241 GTGCTTGGCTTTTATGTGACGCTGTGTGTGAAACCGTGTGTGAGACCACTATGATG 300

QY 405 CCGTGGCCCGCGCTCATGAGCTGTGTGTGCGGGCTTGTGGAAGCAAGAGAGCA 464

Db 301 CCGTGGCCCGCGCTCATGAGCTGTGTGTGCGGGCTTGTGGAAGCAAGAGAGCA 360

QY 465 GCGCGGCTGTGCGGCGCAGCTCATCGCTACGCCCAACCTGGGCAAGTGTCTATCTG 524

Db 361 GCGCGGCTGTACGCGGCGCAGCTCATCGCTACGCCAGGGCTGTGCGGCTGTCTATCTG 420

QY 525 CGCAGCTCAGACCGCAGCTCTACAGCGCTTCCCGACGGCCGACAGCTGTGTGACGA 584

Db 421 CGCTCCCTCAGACCGGCGGTGTGAAACCGCTTCCCGACATGACAGCTGTGTGAGGCT 480

QY 585 GCGTTTATGATCCGCGGAGAAACAGCAAGTGTGAGAAACAGTGTGTACCAACATG 644

Db 481 GGGTTTATGACCGCGGAGAGCGAAGCAAGTGTGAAACCTGATCTATCAACAG 540

QY 645 TTCTGGTGCCTGTGGTGTGTGTTGCCAACCTGTCAATGAAGCGCTGTGTGAGTGA 704

Db 541 TACTGGTGCCTGTGGTGTGTGTTGCCAACCTGTGTGAGCGGCGGAGCGGCGG 600

QY 705 ATCCGGGACCTTATCCGTGCGCAGAGCTGTGTGAGAGAGTGTGTGTGTGTGTG 764

Db 601 ATCCGGGACCAAGCGGCTTAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 660

QY 765 TGTGACACCTGTATGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 824

Db 661 TGTGAGATGCTTCTTCACTATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 720

QY 825 ACTGTGGCGGTACAGCTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 884

Db 721 ACCATGCGCTGTACACACTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 780

QY 885 GCCAAGGCTACCTGTGCGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 944

Db 781 GCTCAGGCTTACAAAGCCAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 840

QY 945 TTCTTCTTATGTTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1004

Db 841 TTCTTCTTACGCGGCTG 900

QY 1005 GATGATGATATTTTGTGACCAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1064

Db 901 GAGATGATGATCTTGTGACCAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 960

QY 1065 GCTGTGATGATGATGACCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1124

Db 961 GCACTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020

QY 1125 CCGAGCCACAGCCCGCTACAGCTGTCTC--CGCCAGTTCCTGAGGCTCTT 1181

Db 1021 GCGGAGGCTGCGGCCCATACAGCGGCTACTGTCTTCAAGCTGCGGAGCTTCTTTC 1080

QY 1182 ATGGGCTCCACTTCAACATCACTGACCTGTGAAAGAGAGATGAGTTCAGC 1233

Db 1081CAGGGCTCACCTTTGACATCACTGCGCCAAAAGAAGACATGCAGTTCAGC 1132

| | RESULT 13 AC051664/C | LOCUS . AC051664 | DEFINITION Homo sapiens chromosome 11 clone RP11-810P12, WORKING DRAFT SEQUENCE, 24 unordered pieces. | ACCESSION AC051664 | VERSION AC051664.2 | KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT. | SOURCE Homo sapiens. | ORGANISM Homo sapiens | REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL |
|--|-------------------------|---------------------|--|-----------------------|-----------------------|---|-------------------------|---|---|
| | AC051664 | 160169 bp | DNA linear HTG 01-SEP-2000 | AC051664 | G1:7637836 | | Homo sapiens. | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | The sequence of Homo sapiens clone Unpublished 2 (bases 1 to 160169) Waterston, R.H. Direct Submission Submitted (15-APR-2000) Genome Sequencing Center, Washington MO 63108, USA On Apr 23, 2000 this sequence version replaced g1:7574984. |

```
misc_feature      86594..36688
                  /note="assembly_name:Contig36"
misc_feature      96789..107169
                  /note="assembly_name:Contig37"
misc_feature      107270..117047
                  /note="assembly_name:Contig38"
misc_feature      117148..127120
                  /note="assembly_name:Contig39"
                  clone_end:SP6
misc_feature      127221..142029
                  /note="assembly_name:Contig40"
misc_feature      142130..160169
                  /note="assembly_name:Contig41"
BASE COUNT      41647 a 36693 c 37296 g 42226 t 2307 others
ORIGIN
Query Match      24.8%; Score 552.8; DB 2; Length 160169;
Best Local Similarity 99.3%; Pred. No. 1.1e-118;
Matches 576; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 1202 CACCTTGACAAAGAGAGAGATGAGTTCAGCCCATCAGAGAGAGAGAGATGCTCA 1261
DB 67594 CACCTTGACAAAGAGAGAGATGAGTTCAGCCCATCAGAGAGAGAGAGATGCTCA 1261
QY 1262 CGCTGGCATATGTCGCGCTTCTAGGCTGAGTCCATGATCAGATCCATCCAGAGG 1321
DB 67535 CGCTGGCATATGTCGCGCTTCTAGGCTGAGTCCATGATCAGATCCATCCAGAGG 1321
QY 1322 AAACCTCAAGGACCAAACTACTGTGGCCCAAGAGAGGATCCCTTCTCCAGAGGCTGCC 1381
DB 67475 AAACCTCAAGGACCAAACTACTGTGGCCCAAGAGAGGATCCCTTCTCCAGAGGCTGCC 1381
QY 1382 CAAAAACCAAGGAGGAGCCAAAGAACTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1441
DB 67416 CAAAAACCAAGGAGGAGCCAAAGAACTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1441
QY 1442 GCTTAAGGCTGTGAGAGCTTCAAGTGTGGCCCAAGTATCAGAGGAGGAGGAGGAGGAG 1501
DB 67356 GCTTAAGGCTGTGAGAGCTTCAAGTGTGGCCCAAGTATCAGAGGAGGAGGAGGAGGAG 1501
QY 1502 TGCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1561
DB 67296 TGCCCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1561
QY 1562 GTCAAGGCTTCAAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1621
DB 67236 GTCAAGGCTTCAAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1621
QY 1622 TTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1681
DB 67176 TTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1681
QY 1682 CCCAGAGATCTCAAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1741
DB 67116 CCCAGAGATCTCAAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1741
QY 1742 GATCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1781
DB 67056 GATCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1781
RESULT 14
LOCUS      BC031186      1956 bp      mRNA      linear      ROD 07-AUG-2002
DEFINITION Mus musculus, clone MGC:37621 IMAGE:4989959, mRNA, complete cds.
ACCESSION   BC031186
VERSION     BC031186.1 GI:21411099
KEYWORDS    MGC.
SOURCE      house mouse.
            Mus musculus
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
REFERENCE      1 (bases 1 to 1956)
AUTHORS      Strausberg, R.
TITLE        Direct Submision
JOURNAL      Submitted (03-JUN-2002) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Jeffrey F. Green, M.D.
CNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mdc@paxil.stanford.edu
          (Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
          R. M.
FEATURES
source
Location/Qualifiers
1..1956
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:37621 IMAGE:4989959"
/tissue_type="Colon, normal, 5 month old male mouse."
/clone_id="NCI CGAP_C024"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
176..1573
/codon_start=1
/product="unknown (protein for MGC:37621)"
/protein_id="AAH31186.1"
/db_xref="GI:21411100"
/db_xref="LocusID:212989"
/transcript="MAL5A1RFLAEBOKRYEFLVYICQVYSLIPVSVLGHVYT
LVAHMMNOYLCPDLMCIYAGTVHGRDRRLYRTLMRAGLSAVILRSVST
AVFRKPTLIDHVRAGFTREKREKLSYNYKVPWCVFSLAIOGREFIDPA
NSALKLLEINVRKSCGMLEFHDWISIPYVOVYIAYSYELACLIQROFIDPA
OGKRDHLDLCVPLFTLLOFFYAGWLKVAQQLINPREDDDDETNLIDNRQVSM
LAVDEMIDDLAMEKDLWDAAEAPFYTAFAFLDQPSFGSTFDIALAKEDMOFO
RLDGVDPDLEHGVDFLQRLPLPAGASVGPAGRRISLRRNSCVSEASTVASCAG
AADGGEVCEGCGDLDLPISLEPELEPACPEPAPIPGPPEPPTVSIQPPAPAP
PWLPSPIGEESPA"
BASE COUNT      411 a 577 c 530 g 438 t
ORIGIN
Query Match      24.2%; Score 538.4; DB 10; Length 1956;
Best Local Similarity 67.6%; Pred. No. 1.4e-115;
Matches 771; Conservative 0; Mismatches 366; Indels 3; Gaps 1;
QY 97 GCTGGCCAGTACCATCTACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 156
DB 39 GCGAGGAGTACCGCTACCTACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 98
QY 157 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 216
DB 99 GCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 158
QY 217 TCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276
DB 159 GTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 218
QY 277 AACAGCTGATTTGAGAAATGACTGCTGATTTGAGAGAGTACATCCAGTCAATCCCA 336
DB 219 AGAAGCGCTACTACAGAGAGCTGTCAATATCTGAGACAGTACGACGAGCTCATCCCG 278
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 10:51:03 ; Search time 6617.31 Seconds
(without alignments)
12237.209 Million cell updates/sec

Title: US-09-622-964-1_COPY_1_5000
Perfect score: 5000
Sequence: 1 ccaaaaattgtctctctg.....ttggagagcgagcgaggt 5000

Scoring table: IDENTITY NUC
Gapop: 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 546 | 10.9 | 2453 | 11 | BC015220 Homo sapi |
| 2 | 541.8 | 10.8 | 545 | 17 | A0588060 CITR1-EI- |
| 3 | 434 | 8.7 | 448 | 9 | A1125626 q994108.x |
| 4 | 433 | 8.7 | 451 | 9 | A1808844 wF97109.x |
| 5 | 431 | 8.6 | 437 | 9 | A1432176 tg77f09.x |
| 6 | 428.8 | 8.6 | 921 | 12 | BE785274 601474631 |

| | | | | | | | |
|----|-------|-----|------|----|----------|-----------|-----------|
| 7 | 424 | 8.5 | 442 | 9 | A1028274 | A1028274 | ov84e04.x |
| 8 | 387.4 | 7.7 | 423 | 9 | A1432177 | A1432177 | tg77f10.x |
| 9 | 296.8 | 5.9 | 2971 | 17 | AF101960 | AF101960 | AF101960 |
| 10 | 293.8 | 5.9 | 2971 | 17 | AF101960 | AF101960 | AF101960 |
| 11 | 272.2 | 5.4 | 1641 | 11 | BC033224 | BC033224 | Homo sapi |
| 12 | 263.8 | 5.3 | 590 | 17 | A0588588 | CITR1-EI- | A0588588 |
| 13 | 263.8 | 5.3 | 814 | 14 | B0440115 | AGENCOURT | AG921904 |
| 14 | 261 | 5.2 | 261 | 9 | AA921904 | cm04f05.8 | AA921904 |
| 15 | 257.6 | 5.2 | 559 | 9 | AU148047 | AU148047 | AU148047 |
| 16 | 257.4 | 5.1 | 652 | 9 | AU120416 | AU120416 | AU120416 |
| 17 | 257.4 | 5.1 | 821 | 12 | BE974661 | BE974661 | BE974661 |
| 18 | 257.2 | 5.1 | 658 | 17 | AG079111 | Pan trogl | AG079111 |
| 19 | 255 | 5.1 | 515 | 14 | B0581726 | 112c10.x | B0581726 |
| 20 | 254.2 | 5.1 | 539 | 17 | A0543001 | RPCI-11-3 | A0543001 |
| 21 | 254.2 | 5.1 | 698 | 14 | B0949414 | AGENCOURT | B0949414 |
| 22 | 253.4 | 5.1 | 698 | 14 | BM999040 | BM999040 | BM999040 |
| 23 | 252.8 | 5.1 | 1038 | 14 | BQ709774 | UI-H-DIO- | BQ709774 |
| 24 | 252.6 | 5.1 | 661 | 17 | AG114592 | Pan trogl | AG114592 |
| 25 | 252.4 | 5.0 | 705 | 17 | AF188083 | UP-479-24 | AF188083 |
| 26 | 250 | 5.0 | 656 | 17 | AG059506 | Pan trogl | AG059506 |
| 27 | 250 | 5.0 | 705 | 17 | AQ346434 | Pan trogl | AQ346434 |
| 28 | 249.6 | 5.0 | 652 | 9 | AU120416 | AU120416 | AU120416 |
| 29 | 248.4 | 5.0 | 481 | 9 | AT791227 | AT791227 | AT791227 |
| 30 | 248.4 | 5.0 | 575 | 9 | AL601708 | AL601708 | AL601708 |
| 31 | 247.8 | 5.0 | 643 | 17 | A2515923 | RPCI-11-4 | A2515923 |
| 32 | 246.2 | 4.9 | 942 | 14 | B0707956 | AGENCOURT | B0707956 |
| 33 | 246 | 4.9 | 676 | 14 | B0181755 | UI-H-EDU- | B0181755 |
| 34 | 245.4 | 4.9 | 654 | 9 | AU117926 | AU117926 | AU117926 |
| 35 | 245.4 | 4.9 | 1358 | 13 | BM564410 | BM564410 | BM564410 |
| 36 | 245 | 4.9 | 889 | 13 | B1457655 | B1457655 | B1457655 |
| 37 | 244 | 4.9 | 867 | 17 | A0782080 | HS-3174.B | A0782080 |
| 38 | 243.8 | 4.9 | 687 | 17 | A0741949 | HS-5569.B | A0741949 |
| 39 | 243.4 | 4.9 | 529 | 12 | BG402924 | AL580585 | BG402924 |
| 40 | 243.4 | 4.9 | 655 | 9 | AL580585 | AL580585 | AL580585 |
| 41 | 243.4 | 4.9 | 778 | 12 | BP663240 | BP663240 | BP663240 |
| 42 | 243.4 | 4.9 | 947 | 12 | BG680848 | BG680848 | BG680848 |
| 43 | 243.4 | 4.9 | 1085 | 13 | BM476343 | BM476343 | BM476343 |
| 44 | 243.2 | 4.9 | 877 | 17 | A0739838 | HS-5505.A | A0739838 |
| 45 | 243 | 4.9 | 494 | 17 | AQ180762 | HS-3220.A | AQ180762 |

ALIGNMENTS

RESULT 1
LOCUS BC015220 2453 bp mRNA
DEFINITION Homo sapiens, clone IMAGE:3877806, mRNA.
ACCESSION BC015220
VERSION BC015220.1 GI:21955361
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2453)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-nhgsc.stanford.edu>
Contact: (Dickson, Mark) mcdexpaxi.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
Series: IRAC Plate: 14 Row: B Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis

This clone has the following problem: Incomplete processing.

FEATURES

SOURCE

Location/Qualifiers

1. .2453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3877806"
/issue_type="lung, large cell carcinoma"
/clone_lib="NIH_MGC_68"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

BASE COUNT 633 a 711 c 606 g 503 t

ORIGIN

Query Match

Best Local Similarity 10.9%; Score 546; DB 11; Length 2453;
Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1037 GCAGGGGCGGACGAGAACGAGACGCTGTGACGACGCGGCTATTCATTCTTCCAP 1096
12 GCAGGGGCGGACGAGAACGAGACGCTGTGACGACGCGGCTATTCATTCTTCCAP 71
1097 GCCCAGAGGGGCTGTCAAGACCCAGGCGCTAGTCAGAGGCTCTCTCTCTGAGAGT 1156
72 GCCCAGAGGGGCTGTCAAGACCCAGGCGCTAGTCAGAGGCTCTCTCTCTGAGAGT 131
1157 CCTGCGACAGAAAGTTGAAGCTCAGCAGACGCGGCTTACCCCACTCTCTCTGCAAGGC 1216
132 CCTGCGACAGAAAGTTGAAGCTCAGCAGACGCGGCTTACCCCACTCTCTCTGCAAGGC 191
1217 TCAGGGGCGGACGAGAACGAGACGCTGTGACGACGCGGCTATTCATTCTTCCAP 1276
192 TCAGGGGCGGACGAGAACGAGACGCTGTGACGACGCGGCTATTCATTCTTCCAP 251
1277 GAGGGGGTGTGCTTCAATTCAGCCCTGTGTCAGCCCAACACCTCTCCAGAGAAAT 1336
252 GAGGGGGTGTGCTTCAATTCAGCCCTGTGTCAGCCCAACACCTCTCCAGAGAAAT 311
1337 TAGAGGGGCGGACGAGAACGAGACGCTGTGACGACGCGGCTATTCATTCTTCCAP 1396
312 TAGAGGGGCGGACGAGAACGAGACGCTGTGACGACGCGGCTATTCATTCTTCCAP 371
1397 TAAAGCAAGAGCTCTTTGAGAGGCTGTGAGGAGTCAAGTGAAGGGGCTCAGC 1456
372 TAAAGCAAGAGCTCTTTGAGAGGCTGTGAGGAGTCAAGTGAAGGGGCTCAGC 431
1457 ACTCAGGTGGGAGTCCAGGCTCTTAAGAGTGGGAGGAGGAGGCTGCGACAGAGTCCAG 1516
432 ACTCAGGTGGGAGTCCAGGCTCTTAAGAGTGGGAGGAGGAGGCTGCGACAGAGTCCAG 491
1517 GAGTCCAGCAGAGCTTATCGCAGAGCTTCTGTGGATCATCGAGCCACCTGGAAGCC 1576
492 GAGTCCAGCAGAGCTTATCGCAGAGCTTCTGTGGATCATCGAGCCACCTGGAAGCC 551
1577 CACCTG 1582
552 CACCTG 557

RESULT 2
A0588060
LOCUS A0588060 545 bp DNA linear GSS 07-JUN-1999
DEFINITION CITR1-El Homo sapiens genomic clone 2643K4, DNA
sequence.
ACCESSION A0588060
VERSION A0588060.1 GI:5014740

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

UNPUBLISHED (1997)

Other GSS: CITR1-El-2643K4.TF

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbeetlgr.org

Clones are available from Research Genetics (Info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.

Seq primer: M13 Reverse

Class: BAC ends.

Location/Qualifiers

1. .545

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="2643K4"

/clone_lib="CITR1-El"

/sex="male"

/cell_type="sperm"

/note="Vector: pBeloBAC11, Site_1: EcoRI; Site_2: EcoRI;

Caltech Human BAC Library D"

BASE COUNT 133 a 151 c 150 g 111 t

ORIGIN

Query Match

Best Local Similarity 10.8%; Score 541.8; DB 17; Length 545;
Matches 543; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

927 TGTCTATTACTAGGGGTGATGAATTCACAGCAACACCATCTTTTCAGATTAAGGGCAC 986
1 TGTCTATTACTAGGGGTGATGAATTCACAGCAACACCATCTTTTCAGATTAAGGGCAC 60
987 TAGGCTGAGAGAGAGAGCTGAATCTACCCGGGGGTACACACACAGGAGGAGGCTGG 1046
61 TAGGCTGAGAGAGAGAGCTGAATCTACCCGGGGGTACACACACAGGAGGAGGCTGG 120
1047 GACCAAGAAACGAGAGCTGTGACTGACGAGCCGGTATTCATTTCTTCATAGCCACAGGG 1106
121 GACCAAGAAACGAGAGCTGTGACTGACGAGCCGGTATTCATTTCTTCATAGCCACAGGG 180
1107 CTGTCAAAAGACCCCAAGGGGCTAGTGAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1166
181 CTGTCAAAAGACCCCAAGGGGCTAGTGAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
1167 AAGTTGAAGCTAGCAGACAGCCCTTAACCCCAACCTCTCTCTCTCTCTCTCTCTCTCTCT 1226
241 AAGTTGAAGCTAGCAGACAGCCCTTAACCCCAACCTCTCTCTCTCTCTCTCTCTCTCTCT 300
1227 GAAACACTGGTGGAGAGATCTTTAGCTCTGATTTTAAAGGCGCATGTAGAGGGGTGT 1286
301 GAAACACTGGTGGAGAGATCTTTAGCTCTGATTTTAAAGGCGCATGTAGAGGGGTGT 360
1287 TGCCCTTAATTCAGAGCCCTGTCTCAGCCCAACACCTCTCAAGAGAAATTAAGAGGGCC 1346
361 TGCCCTTAATTCAGAGCCCTGTCTCAGCCCAACACCTCTCAAGAGAAATTAAGAGGGCC 420
1347 ATGGCCAGGCTGTGCTAGCCGTTGCTTCTGAGCAGATTACAGAGGAGCTAAGACAGG 1406
421 ATGGCCAGGCTGTGCTAGCCGTTGCTTCTGAGCAGATTACAGAGGAGCTAAGACAGG 480

[illegible]

| Query Match | 8.7% | Score 433 | DB 9 | Length 451 |
|-------------|---|---|------|------------|
| Y | 2356 | TTCCAGGTGGATGCTTGGGAAACATCTTTCTTAAGGAAACCTGAGTGGGAAGAACCA | 2415 | |
| Db | 187 | TTCCAGGTGGATGCTTGGGAAACATCTTTCTTAAGGAAACCTGAGTGGGAAGAACCA | 246 | |
| Y | 2416 | TGCAGGTATCTCAGGAAGAGCTTCCTCAGCAGGAAGATCAGCAGGTGGAAGCCCTG | 2475 | |
| Db | 247 | TGCAGGTATCTCAGGAAGAGCTTCCTCAGCAGGAAGATCAGCAGGTGGAAGCCCTG | 306 | |
| Y | 2476 | GAGCCACCATTCAGTAACATCATTTGAGCATCTCTACAGCTAGGTTCCATTATGGGAA | 2535 | |
| Db | 307 | GAGCCACCATTCAGTAACATCATTTGAGCATCTCTACAGCTAGGTTCCATTATGGGAA | 366 | |
| Y | 2536 | TGGAAATATGTTGTGGCAGAGGCTGCGCTGCTCCATCACTTCTCACACTAGGCTGG | 2595 | |
| Db | 367 | TGGAAATATGTTGTGGCAGAGGCTTGCCTGTCTCCCTTTCACTTCTCACACTAGGCTGG | 426 | |
| Y | 2596 | TTGAGAGAGCTTGGAGAGCTAAC 2617 | | |
| Db | 427 | TTGAGAGAGCTTGGAGAGCTAAC 448 | | |
| RESULT 4 | | | | |
| A1808844 | | 451 bp | mRNA | linear |
| LOCUS | wf97h09.x1 Soares_NSF_F8_9W_OT_PA_P_S1 | Homo sapiens | CDNA | clone |
| DEFINITION | IMAGE:2363585 3', mRNA sequence. | | | |
| ACCESSION | A1808844 | | | |
| VERSION | A1808844.1 | GI:5395410 | | |
| KEYWORDS | EST. | | | |
| SOURCE | human. | | | |
| ORGANISM | Homo sapiens | | | |
| REFERENCE | Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| AUTHORS | NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap. | | | |
| NOTE | 1 (bases 1 to 451) | | | |
| JOURNAL | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), | | | |
| COMMENT | Unpublished (1997) | | | |
| | Tumor Gene Index | | | |
| | Contact: Robert Strausberg, Ph.D. | | | |
| | Email: cgaps-remail.nih.gov | | | |
| | This clone is available royalty-free through LNL ; contact the | | | |
| | IMAGE Consortium (info@image.llnl.gov) for further information. | | | |
| | Insert Length: 1234 Std Error: 0.00 | | | |
| | Seq primer: -40UP from Gldco | | | |
| | High quality sequence stop: 444. | | | |
| FEATURES | Location/Qualifiers | | | |
| Source | 1..451 | | | |
| | /organism="Homo sapiens" | | | |
| | /db_xref="taxon:9606" | | | |
| | /clone="IMAGE:2363585" | | | |
| | /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1" | | | |
| | /lab_host="DH10B" | | | |
| | /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with | | | |
| | a modified polylinker; Site_1: Not I; Site_2: Eco RI; | | | |
| | Equal amounts of plasmid DNA from five normalized | | | |
| | libraries were mixed, and ss circles were made in vitro. | | | |
| | Following HAP purification, this DNA was used as tracer in | | | |
| | a subtractive hybridization reaction. The driver was | | | |
| | PCR-amplified cDNAs from pools of 5,000 clones made from | | | |
| | the same 5 libraries. The pools consisted of the following | | | |
| | 1 libraries and clones: Soares NBHSF pool 1: | | | |
| | 309384-310919, 323208-325895 Soares NB2HP pool 1: | | | |
| | 145032-147335, 147720-148103, 148872-149255, 15002 - | | | |
| | 150407, 151176-152327 Soares NB2HR-9W pool 1: | | | |
| | 758280-760583, 772104-774407 Soares NBHP pool 1: | | | |
| | 304776-306311, 320136-322823, 326280-326653 Soares NBHOT | | | |
| | pool 1: 723720-726407, 739080-740999 Subtraction by Bento | | | |
| | Soares and M. Fatima Bonaldo." | | | |
| BASE COUNT | 125 a 90 c 128 g 107 t 1 others | | | |
| ORIGIN | | | | |

Best Local Similarity 98.6%; Pred. No. 2,4e-46; Matches 436; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2176 TCGTCAAGTAATTTTATGAGCGCTTCTAGAGAACACAGAGAGCTTCATCTG 2235
|||||
Db 10 TCGTCAAGTAATTTTATGAGCGCGCTTCTAGAGAACACAGAGAGCTTCATCTG 69
|||||

QY 2236 AGGAGAACAGGAGAGAACAGGAGATATCTGTATTAATTTCAAGTAGTAGTAGTC 2295
|||||
Db 70 AGGAGAACAGGAGAGAACAGGAGATATCTGTATTAATTTCAAGTAGTAGTAGTC 129
|||||

QY 2296 TCTCTAGAAATATCAAGCAAGTGAGAGACACAGACACCGGCGAGAGGCGCTCTAT 2335
|||||
Db 130 TCTCTAGAAATATCAAGCAAGTGAGAGACACAGACACCGGCGAGAGGCGCTCTAT 189
|||||

QY 2356 TTCCAGGTGTGATGTGGGACATCTCTTAAGGGAAGCTGAGTGGGAAGCA 2415
|||||
Db 190 TTCCAGGTGTGATGTGGGACATCTCTTAAGGGAAGCTGAGTGGGAAGCA 249
|||||

QY 2416 TGCAGGTATCTCAGGAAGAGCTTCTCCAGGAGGAAGATCAGAGGTGGAAAGCCCTG 2475
|||||
Db 250 TGCAGGTATCTCAGGAAGAGCTTCTCCAGGAGGAAGATCAGAGGTGGAAAGCCCTG 309
|||||

QY 2476 GAGCCACCATTCACATTAACATCTTGAAGATCTCTACAGCTAGTTCATTTGGGA 2535
|||||
Db 310 GAGCCACCATTCACATTAACATCTTGAAGATCTCTACAGCTAGTTCATTTGGGA 369
|||||

QY 2536 TGGGAATATGTGGTGAGACAGGCTGCTGCTTCCATCTCTACACTAGGTGG 2595
|||||
Db 370 TGGGAATATGTGGTGAGACAGGCTGCTGCTTCCATCTCTACACTAGGTGG 429
|||||

QY 2596 TTGAGAGAGCTTGGAGCTAAC 2617
|||||
Db 430 TTGAGAGAGCTTGGAGCTAAC 451
|||||

RESULT 5
A1432176
LOCUS
DEFINITION
A1432176.437 bp mRNA linear EST 30-MAR-1999
t977f09.x1 Soares_NHMPU.S1 Homo sapiens cDNA clone IMAGE:2114825
3' similar to contains element DBR repetitive element ;, mRNA
sequence.
A1432176
A1432176.1 GI:4308434
EST.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 1153 Std Error: 0.00
Seq primer: -40bp from G100
High quality sequence stop: 401.
Location/Qualifiers
1. 437
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2114825"
/clone_11b="Soares_NHMPU.S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site: 1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three

normalized libraries (melanocyte 2NbhM, pregnant uterus
NbhPU, and fetal heart NbhL19) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT 121 a 89 c 123 g 104 t

ORIGIN

Query Match 8.6%; Score 431; DB 9; Length 437;
Best Local Similarity 100.0%; Pred. No. 4,4e-46;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2176 TCGTCAAGTAATTTTATGAGCGCTTCTAGAGAACACAGAGAGCTTCATCTG 2235
|||||
Db 7 TCGTCAAGTAATTTTATGAGCGCTTCTAGAGAACACAGAGAGCTTCATCTG 66
|||||

QY 2236 AGGAGAACAGGAGAGAACAGGAGATATCTGTATTAATTTCAAGTAGTAGTAGTC 2295
|||||
Db 67 AGGAGAACAGGAGAGAACAGGAGATATCTGTATTAATTTCAAGTAGTAGTAGTC 126
|||||

QY 2296 TCTCTAGAAATATCAAGCAAGTGAGAGACACAGACACCGGCGAGAGGCGCTCTAT 2335
|||||
Db 127 TCTCTAGAAATATCAAGCAAGTGAGAGACACAGACACCGGCGAGAGGCGCTCTAT 186
|||||

QY 2356 TTCCAGGTGTGATGTGGGACATCTCTTAAGGGAAGCTGAGTGGGAAGCA 2415
|||||
Db 187 TTCCAGGTGTGATGTGGGACATCTCTTAAGGGAAGCTGAGTGGGAAGCA 246
|||||

QY 2416 TGCAGGTATCTCAGGAAGAGCTTCTCCAGGAGGAAGATCAGAGGTGGAAAGCCCTG 2475
|||||
Db 247 TGCAGGTATCTCAGGAAGAGCTTCTCCAGGAGGAAGATCAGAGGTGGAAAGCCCTG 306
|||||

QY 2476 GAGCCACCATTCACATTAACATCTTGAAGATCTCTACAGCTAGTTCATTTGGGA 2535
|||||
Db 307 GAGCCACCATTCACATTAACATCTTGAAGATCTCTACAGCTAGTTCATTTGGGA 366
|||||

QY 2536 TGGGAATATGTGGTGAGACAGGCTGCTGCTTCCATCTCTACACTAGGTGG 2595
|||||
Db 367 TGGGAATATGTGGTGAGACAGGCTGCTGCTTCCATCTCTACACTAGGTGG 426
|||||

QY 2596 TTGAGAGAGCT 2606
|||||
Db 427 TTGAGAGAGCT 437
|||||

RESULT 6
BE785274 921 bp mRNA linear EST 20-OCT-2000
LOCUS
DEFINITION
601474631P1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3877806 5',
mRNA sequence.
BE785274
BE785274.1 GI:10206472
EST.
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
NIH-MGC http://mhc.ncbi.nlm.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: DCTD/DP/Genetics
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov

[illegible][illegible]

| QY | Db | QY | Db |
|--|--|---|--|
| 2481 | 311 | 2541 | 371 |
| ACCATTTAGTAAATCATCTATTGACACACTCTACACAGGTAGGTTCCATTATGGGAATGGGA | ACCATTCAGTAAACATCATTTGACACACTCTACACAGTAGGTTCCATTATGGGAATGGGA | ATATGCTGCTGTCGACAGGGCTGCTCTGCTGCTGCTCCATCATCTTCTCACACTAGGGT | ATAAGGTGCTGTCGACAGGGCTGCTCTGCTGCTCCATCATCTTCTCACACTAGGGT |
| 2540 | 370 | 2593 | 423 |

| | |
|------------|---|
| RESULT | 9 |
| AFI01960 | |
| LOCUS | 2971 bp DNA linear GSS 06-NOV-2000 |
| DEFINITION | Human Homo sapiens genomic clone PTM859.14, DNA sequence. |
| ACCESSION | AFI01960 |
| VERSION | AFI01960.1 GI:4193786 |
| KEYWORDS | GSS. |
| SOURCE | Human. |
| ORGANISM | Homo sapiens |

| REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT |
|---------------------|---|---|---------------------------------|---------------------|
| 1 (bases 1 to 2971) | Bepler, G., O'Brian, K. C., Kim, Y. C., Schreiber, G. and Pitterle, D. M. | A 1.4-Mb high-resolution physical map and contig of chromosome segment 11p15.5 and genes in the LOH11a metastasis suppressor region | Genomics 55 (2), 164-175 (1999) | Contact: Bepler, G. |

Duke University Medical Center
Box 2610, MSRB, Room 117, Durham, NC 27710, USA
part of a 1.4 megabase contig including the *LOH1A* metastasis
suppressor region Bin T
Class: unknown.

| FEATURES | source | Location/Qualifiers |
|------------|--------|---------------------------|
| | | 1..2971 |
| | | /organism="Homo sapiens" |
| | | /db_xref="taxon:9606" |
| | | /map="11p15.5" |
| | | /clone="pTWB59.14" |
| | | /clone_11b="Human" |
| | | /note="Vector: pMOS Blue" |
| BASE COUNT | 1121 a | 589 c 514 g 747 t |
| ORIGIN | | |

| Query Match | 5.9% | Score 26.8 | DB 17 | Length 2971 |
|-----------------------|--------------|---|----------------|-------------|
| Best Local Similarity | 68.5% | Pred. No. 2e-29 | | |
| Matches 522 | Conservative | 0 | Mismatches 212 | Indels 28 |
| | | | Gaps | 7 |
| QY | 3974 | TTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACAGGGTTTCACTCCATCAGCCAGGCTGGATGC | 4033 | |
| | | | | |
| Db | 470 | TTATTATTATTATTATTATTATTTTCCGAGACAGTGTCTCACTGTGTTGCCAGACTGGAGTGC | 529 | |
| | | | | |
| QY | 4034 | ACTGGGGGCAATCTTGCTCACTGTAACTCTGCCTCCAGGTGCAGACGATTCCTCCGC | 40933 | |
| | | | | |
| Db | 530 | ACTGGCA-CAACTGGGCTCACTGTGAACCTCGCCTCTGGGGTTCACGCAATCTCTCC | 588 | |
| | | | | |
| QY | 4094 | CTTAGGCCTCCTAGTACGTGGAATATATAGGACAAACACCACAGCGCTGGCTAAATTTTTT | 41533 | |
| | | | | |
| Db | 589 | CTCAGCCTCCCAAGTAGTGGGATTACAGGGCGCCGCCACCAACAACGTAATTTTT--- | 645 | |
| | | | | |
| QY | 4154 | TTTTTCTGTATTTTATAGACAGAGGGTTTCATCATGTGGCCAGGCTGGTCTTAAC | 42133 | |
| | | | | |
| Db | 646 | -----TGTATTTTATAGAGATGAGAGTTTTCACAATGTGGCCAGGCTGGTCTCAAC | 698 | |
| | | | | |
| QY | 4214 | CCCTGACCTCAAGTATCCACCACCTCGGCTCCCAAGTCTGGGATTTACAGGTGCA | 42733 | |
| | | | | |
| Db | 699 | TCCTGACCTC-GTATATCCACCACCTAGCCTCCCAAGTCTGGGATTTACAGGCATAA | 756 | |
| | | | | |
| QY | 4274 | GCACACATGCACAGCCC-ACATGTATACATTTTTTAAATTTATTTTAAATATGTTT | 43322 | |
| | | | | |
| Db | 757 | GCACACGATCTGGGCCCAACATATATATTTTATNGGAAAGCATGAATCTTTCTTAAGATC | 816 | |
| | | | | |

| | | | |
|----|------|---|------|
| QY | 4333 | ATCTAAGCCAGTAGCACTGAGCTGCGCTCTGAAATCCGACAGACTTTGAGGGGCCAAAGTG | 4332 |
| Db | 817 | AGAAACACCCACGGCACAGCTGGCTACGCGCTTAAATCCGCGCACTTTGGGAGGCCGAGCG | 876 |
| QY | 4393 | CGGGGATCAGTTGAGACCTG----GGAGTTACAGCGTGGGCAACATATGATGAGACCCCGCTC | 4448 |
| Db | 877 | GCGAGATACAGAGGTACGAGAGATCGAAGACATCTCCGCGCTAAACAGCGTAAACCCCGTTC | 936 |
| QY | 4449 | TACCAAAATTT---TAAAAATTTAGCTGGGAGTGGTGCGCACTTTCCTGTGGTCCAGCT | 4504 |
| Db | 937 | TACTAAAAATAAAAATTAATAAAATTTAGCTGGGCGTAGTGGTGGGCACTGTAGTCCAGCT | 996 |
| QY | 4505 | ACTTGGGAGAGCTGAGCTGTGGGAGAGCGTGAACCTGTGAGGTGCAAGCGCTCAAGTACCT | 4566 |
| Db | 997 | ACTCGGAGGCTGAGCGCGGAGAAATGGTGTGAACCCAGAGAGCGGAGTGTCCAGTGAACC | 1056 |
| QY | 4565 | ATGATCACACCACTGACTTCAAGCTCGTAGTGACAG-----CTATCTCAAAAGCAACA | 4618 |
| Db | 1057 | AAGATCAATGCCACTGTCACTCCAGCCTTAGGCCAGCAGGGGAGAGCTCCGACTCAAAAAAAA | 1116 |
| QY | 4619 | AAATATATTTATCTAAACGGTAGGATTAATACAGAAATATATGATATGCAATTTAAAT | 4678 |
| Db | 1117 | AAAAATTTAGAAATGAACATGAGATATCTGCTCACTCATTTCTCTTCAACACTGTACTA | 1176 |
| QY | 4679 | GAAGAGAGATTATGATTACATGATTTGAAATTTCAAAAT | 4720 |
| Db | 1177 | GAATATCTATCCAGGCAATCAGGCAAGAAAGAAAAA 1218 | |

| | | | |
|----|------|---|------|
| QY | 4026 | GGAGTGACAGTGGGGGCAATCTTTGGGCACAGTGAACCTCCCTCCACCGAGGACAAGCAT | 4085 |
| Db | 1078 | TGGAGTGACGT-GGCATGATCTTTGGCTCACTGCACAGCTCCGGCTCTTGGGTTACACCAT | 1020 |
| QY | 4086 | TCCTCGCTAGGCTCCTCGTAGTAGCTGGATTATAGGCACACACACACACAGCTGGCTA | 4145 |
| Db | 1019 | TCCTCCCGCTAGGCTCCCGAGTAGCTGGAGCTACAGAGTGGCCACACTGCACCCAGATA | 960 |
| QY | 4146 | ATTTTTTTTTTTTTCTGATTTTATAGTAGACAGGGTTTCATATGTGGCCAGGCTGG | 4205 |
| Db | 959 | ATTTTTTATTTT-----TATTTTAGTAGAAGCAGGGGTTTCCACCGTGTACCCAGGATGG | 905 |
| QY | 4206 | TCTTGAACCCCTGACACCTCAAGTAGATCCACCCACCTCGGGCTCCCAAGTGTGGAGTATC | 4265 |
| Db | 904 | TCTCGATCTCTTGACCTC--GTGATGTGCCCGCTCGGCTCCCAAGTGTCCAGAGATTAC | 847 |
| QY | 4266 | AGGTGTACGCCACCATGCACAGCCACATGGTATCTTTTAAATTAATTTTAAATTA | 4325 |
| Db | 846 | AGGGGTAGGCCACATGTGGCTGGCTTTCTGATCTTTAGAAAACATTCAGTCTTACACAT | 787 |
| QY | 4326 | AATGT-TTAICTAAGGCCAGTAGACAGTGACTCGGCTCTGTATCTCCAGACATTTGAGGGG | 4384 |
| Db | 786 | AAATATATATTTGGGCCAGATGGGGTGGCTTATAGCTGTATCCACACATTTGGGAGG | 727 |
| QY | 4385 | CCAAAGTCCGGGGGACACTTACGCTGG----GAGTTCAGGCTGGGCAACATAGTAGAC | 4440 |
| Db | 726 | CTGAGGTGGGGGATCAGAGAGTAGAGATTGTGAGACACGCTTGGCCAACTTTGTGAAC | 667 |

| | | | | | |
|------------|---|---|---------|--------|-----------------|
| RESULT 10 | AF101960/c | 2971 bp | DNA | linear | GSS 06-NOV-2000 |
| LOCUS | AF101960/c | | | | |
| DEFINITION | AF101960 | Human Homo sapiens genomic clone PTWB59.14, | | | DNA sequence. |
| ACCESSION | AF101960 | | | | |
| VERSION | AF101960.1 | GI:4193786 | | | |
| KEYWORDS | GSS. | | | | |
| SOURCE | human. | | | | |
| ORGANISM | Homo sapiens | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | | | |
| REFERENCE | 1 (bases 1 to 2971) | | | | |
| AUTHORS | Bepier,G., O'Brian,K.C., Kim,Y.C., Schreiber,G. and Pitterle,D.M. | | | | |
| TITLE | A 1.4-kb high-resolution physical map and contig of chromosome | | | | |
| | segment 11p15.5 and genes in the LOH11A metastasis suppressor | | | | |
| | region | | | | |
| JOURNAL | Genomics | 55 (2), | 164-175 | (1999) | |
| MEDLINE | 99134284 | | | | |
| COMMENT | Contact: Bepier G | | | | |

| | LOCUS | BC033224/c | 1641 bp | mrna | linear | HTC 27-JUN-2002 |
|----|-------|--|---------|------|--------|-----------------|
| Db | 666 | CTCATCTCTACTAAAAA-TACAAAAATTAGCTGTTGGTGGCGGGCGGCTCTTAATCCC | | | | 608 |
| Qy | 4501 | AGCTACTCTGGGAACCTAGAGTGTGGGATGGCTGAAGCTGTGAGTGTGAGGCTGCATGTG | | | | 4560 |
| Db | 607 | AGCTACTCTGGAGGCTGAGCAGAGAGATCCCTTGAAACCCAGAGCGGAGGTTGCAGTGG | | | | 548 |
| Qy | 4561 | AGCTATGATGCACACACTGCACCTTCAGGCTGAGTGACAGGCTATCTCAAAAGCAAAACA | | | | 4620 |
| Db | 547 | AGCCAAAGTTTGTGGCAGCTGCACCTCCAGTCTGTGGCAACAGAGTAGACACCTCTCGGAAA | | | | 488 |
| Qy | 4621 | ATAATGTTTATCTAAACGGTAAGTATATATCACGAATATATGAT | | | | 4665 |
| Db | 487 | AATTAATTAATTAATAATAATAATTAACAAAATTATATATATATTAAT | | | | 443 |

```

FEATURES
  source
    Box 2610, MSRB, Room 117, Durham, NC 27710, USA
    part of a 1.4 megabase contig including the LOH11A metastasis
    suppressor region B1n T
    Class: unknown.
    Location/Qualifiers
      1..2971
      /Organism="Homo sapiens"
      /db_xref="taxon:9606"
      /map="11p15.5"
      /clone="PTWB5.14"
      /clone_11b="Human"
      /note="Vector: pMOS Blue"
BASE COUNT      1121 a      589 c      514 g      747 t
ORIGIN

```

| | | |
|-----------|--|------------|
| VERSION | HC033224.1 | GI:2161936 |
| KEYWORDS | HTC. | |
| SOURCE | human. | |
| ORGANISM | Homo sapiens | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | |
| TITLE | 1 (bases 1 to 1641) | |
| JOURNAL | Strausberg, R. | |
| | Direct Submission | |
| | Submitted (25-JUN-2002) National Institutes of Health, Mammalian | |
| | Gene Collection (MGC), Cancer Genomics Office, National Cancer | |
| | Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, | |
| | USA | |
| REMARK | NIH-MGC Project URL: http://mgc.ncl.nih.gov | |

| | | | | |
|-----------------------|--------------|--|----------------|-------------|
| Query Match | 5.9% | Score 293.8 | DB 17 | Length 2971 |
| Best Local Similarity | 69.4% | Pred. No. 4.8e-29 | | |
| Matches 489 | Conservative | 0 | Mismatches 202 | Indels 14 |
| | | | Gaps 6 | |
| QY | 3966 | CATGGTACATTTTTTTTTTTTTTTTATGACAGGCTTCACTCCATCACCAGG | 4022 | |
| | | | | |
| Db | 1138 | CATGTTTCATTTTCATCTTTTTTTTTTTTAGTCGGAGTCTCCGCCGTGCGCTAAGC | 1075 | |

Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Joseph, Jennifer Asano, Ian Rosset, Yaron Butterfield

OY 4143 CTAATTTTTTTTTTTCTGATATTTAGTAGACAGAGGTTTCATCATGTTGGCCAGGC 4202
 DB 211 TTAATTTT-----TGTATTTTAGTAGAGATGGGGTTTATCATGTTGGCCAGGC 162
 OY 4203 TGGCTTTGAACCCCTGACCTCAAGTATCCACCCAGCCCTCCGCTCCCAAAATGCTGGAGT 4262
 DB 161 TGGCTTTGAACCTGACCTCAAGTATCATGTTGGCCAGCCCTCCCAAAATGCTGGAGT 102
 OY 4263 TACAGGTGTGACGACCATGACAGCCCA 4291
 DB 101 TACAGGTGTGACGACCATGCTGCTGGCTCA 73

 RESULT 13
 BQ440115 814 bp mRNA linear EST 24-MAY-2002
 LOCUS AGENCOURT_7826937 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6152819
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ440115 GI:21179191
 VERSION BQ440115
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 814)
 NIH-MGC http://mgi.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M13491 row: e column: 12
 High quality sequence stop: 588.
 FEATURES
 source Location/Qualifiers
 1..814
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6152819"
 /clone_lib="NIH_MGC_67"
 /tissue_type="retinodlastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: PCMV-SPORE6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."
 BASE COUNT 168 a 213 c 183 g 248 t 2 others
 ORIGIN
 Query Match 5.3%; Score 263.8; DB 14; Length 814;
 Best Local Similarity 77.4%; Pred. No. 7e-25; Indels 13; Gaps 3;
 Matches 362; Conservative 0; Mismatches 93;
 OY 3844 AGAGATGAGTCTCACTGTGTGTCAGAGCTGCTCGAACTCTAGGCTCAAGCAATCC 3903
 DB 136 AGAGTGGGGCTTCATATGTTGCCAGAGCTGTTTGAACCTTGGCCTCAAGCAATCC 195
 OY 3904 CCGTGCCTTACGCTCCCAAGGGCTGGATTACAGGTGAGCTACTGCACTTGACCA-- 3961
 DB 196 TCGTCCCTTACGCTCCCAAGGGCTGGATTACAGGTGAGCTACTGCACTTGACCA-- 255
 OY 3962 ----ACCACTGAGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4017
 DB 256 AAGCATTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 315
 OY 4018 ACCGAGCTGAGTGTGAGTGGGGCAATCTTGGCTCACTGTAACTCTCTCCAGAGTG 4077

DB 316 GTCCAGACTGGAGTGGAGT-GGTGTGATCTTGGCTACTGCACCACTTCACTCCAGAGTG 374
 OY 4078 CAAGCATTTCTCCCTTACGCTCTGAGTGTGATTTAGGACACACACACAGC 4137
 DB 375 CACATTAATTTCTGCTCTTATGCTTCTGATGCTGGGGCTATPAGGACATGCGACAGC 434
 OY 4138 CCGGCTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 4197
 DB 435 CCGGATTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 488
 OY 4198 CAGGCTGCTGAGACCCCTGACCTCAAGTATGATTCACACCCAGCCGCGCTCCCAAGTGGT 4257
 DB 489 CAGGCTGCTGAGACCCCTGACCTCAAGTATGATTTTGGCCCTCAGCTCCCAAGTGGT 548
 OY 4258 GGGATTACAGGTGTGACGACCATGACAGCCACATGATGATTTT 4305
 DB 549 GGGATTACAGGCGGTGAGCGACATGTCGACAGGCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 596

 RESULT 14
 AA921904 261 bp mRNA linear EST 23-JUN-1998
 LOCUS omd0f05.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
 DEFINITION IMAGE:1543521 3', mRNA sequence.
 ACCESSION AA921904
 VERSION AA921904.1 GI:3069213
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 261)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 365 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham.
 FEATURES
 source Location/Qualifiers
 1..261
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1543521"
 /clone_lib="Soares_NFL_T_GBC_S1"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: pRTT3D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NBH119w, testis NHT, and B-cell
 NCI-CGAP-GCB1) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."
 BASE COUNT 81 a 48 c 81 g 51 t
 ORIGIN
 Query Match 5.2%; Score 261; DB 9; Length 261;
 Best Local Similarity 100.0%; Pred. No. 3.1e-24;
 Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2211 GAACACAAAGAGAGCTTCATCTGTGAGGAAAGAGGAGGAGGAGGAGATATCTGT 2270
 DB 1 GAACACAAAGAGAGCTTCATCTGTGAGGAAAGAGGAGGAGGAGGAGATATCTGT 60

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 18:10:38 ; Search time 654.38 Seconds
(without alignments)
10693.617 Million cell updates/second

```
Title: US-09-622-964-1_COPY_1_5000
Perfect score: 5000
Sequence: 1 ccaaaaatgtctcttg.....ttggagagccgagcgsgt 5000
```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

```
Minimum DB seq length: 0
```

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Database : Published_Applications_NA:*

```

1: /cgn2_6/p/odata/2/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/p/odata/2/pubna/PCT_NEW_PUB.seq.*
3: /cgn2_6/p/odata/2/pubna/US06_NEW_PUB.seq.*
4: /cgn2_6/p/odata/2/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/p/odata/2/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/p/odata/2/pubna/PCTOS_PUBCOMB.seq.*
7: /cgn2_6/p/odata/2/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/p/odata/2/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/p/odata/2/pubna/US09_NEW_PUB.seq.*
10: /cgn2_6/p/odata/2/pubna/US09_PUBCOMB.seq.*
11: /cgn2_6/p/odata/2/pubna/US10_NEW_PUB.seq.*
12: /cgn2_6/p/odata/2/pubna/US10_PUBCOMB.seq.*
13: /cgn2_6/p/odata/2/pubna/US06_NEW_PUB.seq.*
14: /cgn2_6/p/odata/2/pubna/US06_PUBCOMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Query Match | Length | DB | ID | Description |
|--------|-----|-------|-------------|--------|----|--------------------|--------------------|
| c | 1 | 375.6 | 7.5 | 17792 | 9 | US-10-091-504-1599 | Sequence 1599, App |
| c | 2 | 375.6 | 7.5 | 17792 | 10 | US-09-764-869-1599 | Sequence 1599, App |
| c | 3 | 350.6 | 7.0 | 13819 | 10 | US-09-764-877-2596 | Sequence 2596, App |
| c | 4 | 340.6 | 6.8 | 13821 | 10 | US-09-764-877-2595 | Sequence 2595, App |
| c | 5 | 336.6 | 6.7 | 4360 | 9 | US-10-161-803-34 | Sequence 34, App |
| c | 6 | 336.6 | 6.7 | 10139 | 9 | US-10-091-483-335 | Sequence 335, App |
| c | 7 | 336.6 | 6.7 | 10139 | 10 | US-09-764-846-335 | Sequence 335, App |
| c | 8 | 334.6 | 6.7 | 14581 | 9 | US-10-216-373-4 | Sequence 4, App1 |
| c | 9 | 334.6 | 6.7 | 22484 | 10 | US-09-875-114-2 | Sequence 2, App1 |
| c | 10 | 334.6 | 6.7 | 22484 | 10 | US-09-880-107-3341 | Sequence 3341, App |
| c | 11 | 331.6 | 6.6 | 32194 | 9 | US-09-764-891-7038 | Sequence 7028, App |
| c | 12 | 329 | 6.6 | 10139 | 9 | US-10-091-483-335 | Sequence 335, App |
| c | 13 | 329 | 6.6 | 10139 | 10 | US-09-764-846-335 | Sequence 335, App |
| c | 14 | 328 | 6.6 | 14448 | 9 | US-09-860-670-250 | Sequence 250, App |
| c | 15 | 328 | 6.6 | 14451 | 9 | US-09-860-670-253 | Sequence 253, App |
| c | 16 | 326.4 | 6.5 | 14426 | 9 | US-09-860-670-252 | Sequence 252, App |
| c | 17 | 321.6 | 6.4 | 8369 | 9 | US-10-091-391-67 | Sequence 67, App1 |
| c | 18 | 321.6 | 6.4 | 8369 | 10 | US-09-764-990-67 | Sequence 67, App1 |
| c | 19 | 321.2 | 6.4 | 5159 | 10 | US-09-764-877-3707 | Sequence 3707, App |

ALIGNMENTS

| | | | | | | | |
|---|----|-------|-----|-------|----|---------------------|--------------------|
| C | 20 | 319.4 | 6.4 | 17397 | 9 | US-10-091-5504-1945 | Sequence 1945, Ap |
| C | 21 | 319.9 | 6.4 | 17397 | 10 | US-09-764-869-1945 | Sequence 1945, Ap |
| C | 22 | 319.4 | 6.4 | 19345 | 9 | US-10-091-5504-1944 | Sequence 1944, Ap |
| C | 23 | 319.4 | 6.4 | 19345 | 10 | US-09-764-869-1944 | Sequence 1944, Ap |
| C | 24 | 318.8 | 6.4 | 7734 | 9 | US-09-764-868-1317 | Sequence 1317, Ap |
| C | 25 | 318.8 | 6.4 | 10093 | 9 | US-10-091-5504-1390 | Sequence 1390, Ap |
| C | 26 | 318.8 | 6.4 | 10093 | 10 | US-09-764-869-1390 | Sequence 1390, Ap |
| C | 27 | 318 | 6.4 | 14426 | 9 | US-09-860-670-249 | Sequence 249, Ap |
| C | 28 | 317.8 | 6.4 | 19334 | 9 | US-10-091-5504-1943 | Sequence 1943, Ap |
| C | 29 | 317.8 | 6.4 | 19334 | 10 | US-09-764-869-1943 | Sequence 1943, Ap |
| C | 30 | 317.6 | 6.4 | 14417 | 9 | US-09-860-670-251 | Sequence 251, Ap |
| C | 31 | 317 | 6.3 | 3471 | 9 | US-10-092-154-1893 | Sequence 1893, Ap |
| C | 32 | 317 | 6.3 | 3471 | 10 | US-09-764-847-1893 | Sequence 1893, Ap |
| C | 33 | 316.6 | 6.3 | 29449 | 9 | US-09-989-442-161 | Sequence 161, Ap |
| C | 34 | 316.6 | 6.3 | 29449 | 9 | US-10-074-045-73 | Sequence 73, Ap |
| C | 35 | 315.4 | 6.3 | 30175 | 10 | US-09-738-878-3 | Sequence 3, Ap |
| C | 36 | 315.4 | 6.3 | 30175 | 12 | US-10-163-381-3 | Sequence 3, Ap |
| C | 37 | 315 | 6.3 | 32189 | 9 | US-09-764-891-8604 | Sequence 8604, Ap |
| C | 38 | 314.8 | 6.3 | 32170 | 9 | US-10-074-093-1108 | Sequence 1108, Ap |
| C | 39 | 314.4 | 6.3 | 32195 | 10 | US-09-764-860-1108 | Sequence 1108, Ap |
| C | 40 | 313.8 | 6.3 | 32195 | 9 | US-09-764-891-6688 | Sequence 6688, Ap |
| C | 41 | 312.8 | 6.3 | 32169 | 9 | US-09-764-891-8605 | Sequence 8605, Ap |
| C | 42 | 312.6 | 6.3 | 14040 | 9 | US-09-764-891-5478 | Sequence 5478, Ap |
| C | 43 | 312.6 | 6.3 | 14040 | 9 | US-09-764-891-10205 | Sequence 10205, Ap |
| C | 44 | 312.4 | 6.2 | 5159 | 10 | US-09-764-877-3707 | Sequence 3707, Ap |
| C | 45 | 312 | 6.2 | 10855 | 9 | US-09-764-891-7965 | Sequence 7965, Ap |

RESULT 1
US-10-091-504-1599/c
Sequence 1599, Application US/10091504

```

APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007C1
CURRENT APPLICATION NUMBER: US/10/091.504
CURRENT FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 2442
Prior Application removed - See File Wrapper or Palm
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1599
LENGTH: 17792
TYPE: DNA
ORGANISM: Homo sapiens
US-10-091-504-1599

```

| | | | | |
|---------------------------|--------|--------------------|------------|---------------|
| Query March | 7.5%; | Score 375.6; | DB 9; | Length 17792; |
| Best Local Similarity | 64.3%; | Pred. No. 1.7e-93; | | |
| Matches 748; Conservative | 0; | Mismatches 359; | Indels 57; | Gaps 10; |

| | | | |
|----|------|--|------|
| QY | 3843 | GAGAGATGAGATCAGTCAATGGTTCTCCAGGCTGGCTGATCTCTAGGCTCAACATTC | 3902 |
| Db | 8715 | GGGAGACGGGGTTTCACCAATGTAGCCAGATGTGTCTGGATCTCTGACATCATC - ATC | 8653 |
| QY | 3903 | CCCTTGCCTTAGCTCCCAAGGGCTGGGATTTACAGTGTGAGCTACTGTCACTTACCCA | 3965 |
| Db | 8657 | CGCCCTCTTGGCTCCCAAGTCTCTGGATTTACAGGATGAGCCACCGCGC - - - - - | 8606 |
| QY | 3963 | CCACATGCTACTTTTTTTTTTTTTTTTTTTTGTAGACAGAGGTTTCACTCATCAACCA | 4022 |
| Db | 8605 | CCAGCTTTCTCTTTCTTTCTTTTTTTTTTTTTTGAACAAGATGTCTGCTCTGTGGCCA | 8544 |
| QY | 4023 | GGCTGGATGTCAGTGGGGGCAATCTTGCTACTGTAACTCTGCTCTCCAGGTGCAGC | 4083 |
| Db | 8545 | GGCTGGATGTCAGT - GGGCGATCTTGGCTTACTGCACCACTCCACCTCCAGTTCAAGC | 8483 |
| QY | 4083 | GATCTCTGCTTAGGCTCTCTGATGACTGGATTAATAGGCACACACACAGCTCTGG | 4144 |
| Db | 8486 | AATTCCTCTGCTCAGCTCTGGAGTACCTAGATTACAGGTGCACACACACACACGCCAG | 8422 |

| | | | |
|--|------|--|------|
| QY | 4143 | CTAATTTTTTTTTTTTTCTATATTTTATAGAGACAGGGGTTTCATATCTTTGGCCAGGC | 4202 |
| Db | 8426 | CTAATGTTT-----TGTATTTTATAGTAAGAAATTTGGGTTTCAACCTGTGGCCAGTC | 8376 |
| QY | 4203 | TGTCCTTGAACCCCTGACCTCAAGTGTCCACCACCTCGGCTCCCAAGATGCTGGAT | 4262 |
| Db | 8375 | TGTCCTGGAACCTCGTACCTCAGGTATTTGCCACCTCGGCTCCCAAGATGCTGGAT | 8316 |
| QY | 4263 | TACAGGCTAGGCCACCATGCACAGCCCACTAGTACATTTTAAATATATTTTAAAT | 4322 |
| Db | 8315 | TACAGGCTAGTCACACGCTGGCTACTGTCTAAATTTTAAAA----- | 8269 |
| QY | 4323 | TAAATGTTTTATCTAAGCCAGTAGACGATGCTCGCTGTGTATCCAGACATTTGAGG | 4382 |
| Db | 8268 | -----GAGCGCGGGGCGAGTGGCTCAGCCCTGTATTCACAGACTTTGGGA | 8223 |
| QY | 4383 | GGCCAGAGTGGGGGATCACTTAGCCTGGGAGTTCAAGCTG--GGCAATATGTGAGA | 4439 |
| Db | 8222 | GGTCGAGATGGGTGGATCACAGGCTCAGAGTTTGAGACCAAGCTCGATATGTGAAA | 8163 |
| QY | 4440 | CCCCGTCTACCAAAATTTAAAAATTTAGCTGGAGTGGTGGCATTTGGCTGTGGTC | 4499 |
| Db | 8162 | ACCGGTCTTAGTAAAAA--TACAAAAATTTAGCCGGGCTGGTGGCAGGACCTTAGTCC | 8104 |
| QY | 4500 | CAGCTACTTGGGAAGCTGAGGTGTGGGGATGGCTGAAGCCTGTGAGGTGAGGCTGAGT | 4559 |
| Db | 8103 | CAGCTACTCAGAGAGCTGAGGACGAGAGATGCTTGAACCTGGGAGGGGAGGTTCAGT | 8044 |
| QY | 4560 | GAGCTATGATCACACCACTGCATCTTCAGCCTG--AGTAGACAGGCTATCTCAAAAGCAAC | 4617 |
| Db | 8043 | GAGCTGAGATTGCACCACTGCATCTTCAGCCTGGGCAACAAAGCGAGACTCCGTCAAAA | 7984 |
| QY | 4618 | AAATATATGTTTATCTAAACGGTAAAGGTATATCACAGAAATATATGATGATTTAAAT | 4677 |
| Db | 7983 | AAAAAATAATATATATATATATCAAAAAATTTAGCCGGCATGTGATGCACCCATATA | 7924 |
| QY | 4678 | TGAAAAAGCATTAATGATTTACATGATGTGTAAAAATATCAAAATACATGAATTCGTGTT | 4737 |
| Db | 7923 | ATCCAGAGTACTCGGAGAGCTGAGGACGAGAAATCGTTGAACCTAGAGAGGTGAAGTTG | 7864 |
| QY | 4738 | C--TTAATTAATGCTAGCAACAAGGCACATTTGGTTTTACTAGGGCAACCAAGTACTTA | 4795 |
| Db | 7863 | CAGTTCTCGAGATCTCAACACTCAGCTCCACGCTGGGCAACAGAGCGAGATTCATCTC | 7804 |
| QY | 4796 | AAAAAAGTTAGGCGCAGCCACAGGAGGGGCTCACACCTGTATATCCAGCAATTTGGAGGCCA | 4855 |
| Db | 7803 | AAAAAAGGCGCAGGTTGGTGGCTCAGACGCTGTATATCCAGCAATTTGGAGGCCA | 7744 |
| QY | 4856 | AGCAGAGAGATCACTTGAGCCCAAGAGATTAGCACTGAGCAACATATGAGAGATCTGA | 4915 |
| Db | 7743 | AGGCAAGGTGATTCCTGTGATCTCAGAGATTCGAGAAC--AGCTGGGCCAATGATGTA | 7686 |
| QY | 4916 | TCTTGTCTCTATAAAAAATTTAAAAAATTTGGCTAGGCCCTTTGGCTTAACCCGTAAATCC | 4975 |
| Db | 7685 | CCCGGTCTCTATAAAAATACAAAAATAGCCGGGTGCTTGGGCTGTACTTTAATCTT | 7628 |
| QY | 4976 | AGCATTTTGGAGGCGCAGCGGG 4999 | |
| Db | 7625 | AGCTTTTCAGAGGCTGAGCGAGG 7602 | |
| RESULT 2 | | | |
| US-09-764-869/c | | | |
| Sequence 1599, Application US/09764869 | | | |
| Patent No. US20020061521A1 | | | |
| GENERAL INFORMATION: | | | |
| APPLICANT: Rosen et al. | | | |
| TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies | | | |
| FILE REFERENCE: PC007 | | | |
| CURRENT APPLICATION NUMBER: US/09/764, 869 | | | |
| CURRENT FILING DATE: 2001-01-17 | | | |
| Prior application data removed - refer to PALM or file wrapper | | | |

[illegible]

| | | | |
|----|------|---|------|
| OY | 4084 | ATTTCACGCGCTTACCGTCGAGTACGCGAAATATATGCGACACACACACCGCTCGC | 4143 |
| Db | 2386 | ATTCCCTCGCCTCACCTCTCGTAGCTGGGACTACACGGCATGTGCGACCAATGCCACGC | 2337 |
| OY | 4144 | TAATTTTTTTTTTTCTGTATTTTATAGTAGACAGGGTTTCATCATGTGGCCAGCT | 4203 |
| Db | 2336 | TAATTTT-----TGATCTTTGGTAGAGATGAGAGTTTACACCATGTGGCCAGCT | 2288 |
| OY | 4204 | GSTCTTGAACCCCTGACCTCAATGATCCACCACCTGGCCCTCCCAAAGTCTGGAGTT | 4263 |
| Db | 2285 | GGCTCTGAACCTCTGACCTCAAGTATGATCACCCACCTCGGCCCTCCCAAAGTCTGGAGTT | 2228 |
| OY | 4284 | ACAGGTGACGACACATGACACAGCCCAATGTGATCTTTTAAATTTATTTTAAT | 4332 |
| Db | 2225 | ATAGGATGAGCACTGTGCCACGCTGAATGTATTAATTTT-----AATTTTTTTAAAGTT | 2171 |
| OY | 4324 | AAATGTTTATCTAAGGCCAGTAGACAGTACGCTCGCTGTATATCCAGCACTTTGAGGG | 4389 |
| Db | 2170 | TTTAATGACAGTAGTAGGCGCGCGGTACTCACTTGTATATCCAAACTTTGGGAG | 2111 |
| OY | 4384 | GCCAAAGTGGCGGGATCACTTGAGCTGGAGTT-----CAGCGTGGCAACATAGTA | 4437 |
| Db | 2110 | GCCAAAGGAGCGGATCACTGAGGTGAGGATTTGAGACCAACCTGTGCCAAATGTGTA | 2051 |
| OY | 4438 | GACCCGCTCTACCAAAAATTTAAAAATATAGCGGGGAGTGTGGCATTTGGCTGTGCT | 4497 |
| Db | 2050 | -AACCATCTCTACTAAAAATATATAAAAATTAGCTGGGCAATGTGTGGCGCTGTAT | 1992 |
| OY | 4498 | CCAGGCTACTTGGGAGCTGAGGTGTGGGATGAGCTGAAGCCTGTGAGGTGAGAGCTCA | 4557 |
| Db | 1991 | CCGAGCTACTTGGGAGGCTGTGGCAGAGATTCGTTTAACATGGGAGCGGAGGCTGCA | 1932 |
| OY | 4558 | GTGAGCTTGTATCACACCACCTGACCTTGAAGCTGTGAGCTATCTCAAAAAGCAAA | 4617 |
| Db | 1931 | GTGAAGCTGAGTTGGCCACTGCACTCCAGTTTGGGCGA-----TAAAGCAAAAATCTGT | 1877 |
| OY | 4618 | AAATATATGTTATCTAAGGCTAAGGTATATATACAGAAATATATGATAGCATTTTAAT | 4677 |
| Db | 1876 | TCTCAAAAAAATTAATTAATTTCAATTATGAAAAAAGAAATTTAAAAACACACATTT | 1817 |
| OY | 4678 | TGAAAAAGCATTAATGATTACATGATGTATAAATATCAATATACATGAATCTTGTGTT | 4737 |
| Db | 1816 | ACAAATGCGATCATTTCCCTCACTCACTCAAAAAAAGAGTTTACTGAGTG | 1757 |
| OY | 4738 | CTTAATAATGCTAGCAACAGGACATTTGTTTTTACTAGGCAACAGGTACTTTAAA | 4797 |
| Db | 1756 | CTAAGTCGTGTT-----CTAGAGTTTAAAGAT | 1730 |
| OY | 4798 | AAAAGTTAAGGCCACGACAGGGGCTCACTGATATCCAGACATTTGGAGGCCAAG | 4857 |
| Db | 1729 | TCCAAAAATAGGCCACGACAGTGGCTCACTGTATATTCAGACATTTGGGAGGCATG | 1670 |
| OY | 4858 | GCAGGAGGATCACTTGAAGCCAGAGATTATAGACCTGTGCAACATATGGGAATCTGTAT | 4917 |
| Db | 1669 | ATGGGTGGATCACTTGAAGGCCAAGAGTTTCGAGATCGTCTGGCCA--CATGTAGAACC | 1612 |
| OY | 4918 | TTGTCTCTATAAAAAATTTAAATTTGGCTATGAGCCCTTTGGCTTACACCCGTAAATCCAG | 4977 |
| Db | 1611 | CTGTCTCTACTAGAAATATCAAAAAATATAGCTGGCATGGTGGCGAGGTGCAATCCAG | 1552 |
| OY | 4978 | CACTTTGGGAGCCGAGGC | 4996 |
| Db | 1551 | CTACTCGGATGCTGAAGC | 1533 |

RESULT 5
 US-10-161-803-34
 ; Sequence 34, Application US/10161803
 ; Publication NO. US20030092028A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ma, Yuanhong
 ; APPLICANT: Lin, Chih-Jian

```

: APPLICANT: Chen, Fan
: APPLICANT: Fairman, Jeffery
: APPLICANT: Chen, Yil-Der I.
: TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS AND
: TITLE OF INVENTION: TREATMENT OF INSULIN RESISTANCE AND RELATED CONDITIONS
: FILE REFERENCE: 421452000300
: CURRENT APPLICATION NUMBER: US/10/161,803
: CURRENT FILING DATE: 2002-06-03
: PRIOR APPLICATION NUMBER: US 60/295,264
: PRIOR FILING DATE: 2001-06-01
: NUMBER OF SEQ ID NOS: 61
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 34
: LENGTH: 4360
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-161-803-34

```

| | | | | |
|-----------------------|-----------------|--------------------|------------|--------------|
| Query Match | 6.7% | Score 336.6; | DB 9; | Length 4360; |
| Best Local Similarity | 71.2% | Pred. No. 5.3e-83; | | |
| Matches 546; | Conservative 0; | Mismatches 179; | Indels 42; | Gaps 6 |

| | | | |
|----|------|--|------|
| QY | 384 | AGAGATGAGCTCAGCTAGTGTGTGTTCACAGCGCTGGCTCGAACTCCTAGGCTCAAGCAATCC | 3903 |
| Db | 1240 | AGAAACAAGGGTTTACCGTATTTGGCAGAGGGCTGGCTTGAATCTCTGACCTGGGGATCT | 1299 |
| QY | 3904 | CCCTGCGCTTACCTCCCAAGAGGGGTGGGATTACAGAGTGTAGCTTACTGACCTTG----- | 3957 |
| Db | 1300 | ACCCTCTGTGGCTCCCAAAATGCTAGATTTACAGGATGAGCCACGCTGCTGGCCCTA | 1355 |
| QY | 3958 | ---ACCAACACATGATCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGAGACAGGGTTTCAC | 4014 |
| Db | 1360 | ATTACCCCAATCTTTTTTTTGT | 1419 |
| QY | 4015 | ATCACCCAGGCTGGAGTGCAGCTGGGGGCAATCTGGCTCAGTGTAACTCTGCTCCACG | 4074 |
| Db | 1420 | GTCGCCCAAGGTGGAGTGCAGT--GGCTGATCTTGGCTCAGCTGCAATGCAAGTCTCGCTCCGG | 1478 |
| QY | 4075 | GTCGAACGATTTCTCTGCGCTTACGCTCTAGTACGTGGAATTATATAGCACACACACC | 4133 |
| Db | 1479 | GTTACGCGCATTTCTCTGCTCTGACCTCTCCAGATCTCCAGATGCTGGGACTRACAGGCCCCACAA | 1538 |
| QY | 4135 | ACGCGCTGGCTAATTTTTTTTTTTTCTGTATTTTGTAGTAGAGACAGGGTTTCATCATGTT | 4194 |
| Db | 1539 | ACGCCAGTTAA-----TTTGTGTATTTTGTAGTAGAGACGCGGTTTTCACGCTGT | 1589 |
| QY | 4195 | GGCCAGGCTGCTCTTGAACCCCTACTCTCAAGTATATCCACCACCTGGGCTCCCAAGT | 4253 |
| Db | 1590 | AGCCAGGATGGTCTCGATCTTCTTACCTCA--TATATGGCCACCTCTGCTCCCAAGT | 1647 |
| QY | 4255 | GCTGGGATTCAGAGTGTACAGCCACATGACACAGCCCAATGTAATTTTTTAAATATAT | 4314 |
| Db | 1648 | GCTGGGATTCAGAGCGTGAAGCCACCGCTGGCCGCCCTA-----AT | 1688 |
| QY | 4315 | TTTTTAATTAATATGTTATCTAAGGCCAGTAGACGATGATCGGCTGTATATCCACAGA | 4374 |
| Db | 1689 | TACCCAGTCTTAAAGAAAAATGGCGGCGGCGGAGTGGCTGAACCTGTAAATCCACAGA | 1748 |
| QY | 4375 | CTTTGAGGGGCGCAAGGTGCGGGGATTCAC--TTGAGCTGGGATTCACGCTGGGCACAT | 4432 |
| Db | 1749 | CTTTGGGAGGCGCGGAGGGGTGTATCAGATCAGAGATCGAGATCCATCTCGGCTTAAAC | 1808 |
| QY | 4433 | AGTGAGACCCCGTCTTACCAAAAAATTTAAAAAATTTAGCTGGAGTGTGGCAATTTGCT | 4492 |
| Db | 1809 | GGTAAACCAACGCTCTCTACTATAAAATATCAAAAAATTTAGCCGGCGTGTGGCGGCGCT | 1868 |
| QY | 4493 | GTCGTCACACTATTGGGAGGCTGAGGTGTGGGATGTGGTGAAGCCCTGTGAGGCTCGAG | 4552 |
| Db | 1869 | GTAATGCCACACTACTCGGAGGCGGAGGACAGAGAAATAGCGGTGAACCTGGGAGGCACAGC | 1928 |
| QY | 4553 | CTGCACTGAGACTATGATCAACACACTGCTACTTACAGCTGAGTGCAG | 4599 |
| Db | 1929 | TTTCAGTAGACTGATGATCTTCAACACTGCTACTTCCAGCTTGGGCGCAG | 1975 |

Db 3582 TCCGGAACCTGGAGGCTGAGGAGAGAAATGCTGAACCTGGAGGAGGAGGCTGCA 3641
QY 4558 GTGAGCTATGATCACAACCACTGACCTTACCTGAGTGACAG 4600
Db 3642 GTGAGCAAGAGACTGCTCATTTGCACTCCACCTGGGCAACAG 3684
RESULT 8
US-10-216-373-4
; Sequence 4, Application US/10216373
; Publication No. US20030096750A1
; GENERAL INFORMATION:
; APPLICANT: Tombran-Tink, Joyce
; APPLICANT: Steele, Fintan R
; APPLICANT: Chader, Gerald J
; APPLICANT: Becerra, Sofia P
; APPLICANT: Johnson, Lincoln V
; APPLICANT: Rodriguez, Ignacio R
; TITLE OF INVENTION: RETINAL PIGMENTED EPITHELIAL DERIVED NEUROTROPIC FACTOR
; FILE REFERENCE: 2026-4203051
; CURRENT APPLICATION NUMBER: US/10/216, 373
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US/08/520, 373
; PRIOR FILING DATE: 1995-08-29
; PRIOR APPLICATION NUMBER: 08/377, 710
; PRIOR FILING DATE: 1995-01-25
; PRIOR APPLICATION NUMBER: 08/279, 979
; PRIOR FILING DATE: 1994-07-25
; PRIOR APPLICATION NUMBER: 07/894, 215
; PRIOR FILING DATE: 1992-06-04
; PRIOR APPLICATION NUMBER: 07/952, 796
; PRIOR FILING DATE: 1992-09-24
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 14581
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: mRNA: 6683; EXON: 6683-6790; EXON 11584-11675;
; OTHER INFORMATION: EXON: 14539-14581; INTRON: 6791-11583; INTRON:
; OTHER INFORMATION: 11676-14538; CDS: 11584-11675; 14539-14580
US-10-216-373-4
Query Match 6.7%; Score 334.6; DB 9; Length 14581;
Best Local Similarity 71.2%; Pred. No. 4.6e-82;
Matches 549; Conservative 0; Mismatches 169; Indels 53; Gaps 6;
QY 3844 AGAGATGAGTCTGACGTGTGTCGAGGCTGTCGAGACTCTCTAGGCTCAACGATCC 3903
Db 13345 AGAGATGAGGCTGACATGTTCTCTCAGTTGGTCTAAACTCTGAGCTCAAGTATCC 13404
QY 3904 CCCTGCTTATGCTCCCAAGGAGGCTGAGATTACAGAGTGAGTACTGACATTGACCAAC 3963
Db 13405 ATCTTCTGCGCTGCAAGAGTGTGAGATTATAGGATTAAGCAACGACGACCTGCTCC 13464
QY 3964 CACAT-----GGTACTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGGGTTTCACT 4012
Db 13465 AATTTTATATTTATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 13524
QY 4013 CCATCAACCCAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4072
Db 13525 CTGTCAACCCAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 13583
QY 4073 AGGTGCAAGGATTTCTCTGCTTACCTTACCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4132
Db 13584 GGGTTTCAAGGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 13643
QY 4133 CCACGCTGCTATATTTTTTTTTTTTTTTTTTTTTTTTGTATTTTATTTAGTATTTAGTAT 4192
Db 13644 CCATGCCCCCTTATATTTT-----TGTATTTTATTTAGTATTTAGTATTTAGTATTT 13694

QY 4193 TTGGCAGGCTGCTTGAACCCCTGACCTCACTGATTCACCCACCTGCGCTGCCAA 4252
Db 13695 TTGGCAGGATGCTGCTGACCTGCTGACCTGCTGACCTGCTGACCTGCTGACCTGCTG 13754
QY 4253 GTGCTGGATTTACAGGTGTACAGCCACATGACAGCCACATGTTATTTTAAATTT 4312
Db 13755 GTGCTGGATTTATGAGTGTGAGCCATGCGCTGA----- 13788
QY 4313 ATTTTATTTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 4372
Db 13789 -TGTATTTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 13846
QY 4373 CACTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4428
Db 13847 CACTTGTGTAGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 13906
QY 4429 ACATATGTAGACCCGCTCTTACCAAAATTTTAAATTTATTTATTTATTTATTTATTTAT 4488
Db 13907 ACAGGAGTAAACCCGCTCTTACCAAAATTTTAAATTTATTTATTTATTTATTTATTTAT 13966
QY 4489 GCGTGTGCTCCACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4548
Db 13967 GCGTGTGCTCCACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 14026
QY 4549 GAGGCTGAGTGTATGATCAGACCACTGACCTTACCTGAGTACAG 4599
Db 14027 GAGCTTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 14077
RESULT 9
US-09-875-114-2
; Sequence 2, Application US/09875114
; Patent No. US20020002131A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020002131A1thwestern University
; APPLICANT: No. US20020002131A1 Bouck
; APPLICANT: Paul Giliis
; APPLICANT: David Dawson
; TITLE OF INVENTION: Methods and Compositions for Inhibiting Angiogenesis
; FILE REFERENCE: 0290-2302
; CURRENT APPLICATION NUMBER: US/09/875, 114
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 09/122, 079
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: PCT/US98/15228
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: US 08/899, 304
; PRIOR FILING DATE: 1997-07-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 22484
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: 1..22484
; OTHER INFORMATION: "n" means either a, c, t, or g
US-09-875-114-2
Query Match 6.7%; Score 334.6; DB 10; Length 22484;
Best Local Similarity 71.2%; Pred. No. 6.3e-82;
Matches 549; Conservative 0; Mismatches 169; Indels 53; Gaps 6;
QY 3844 AGAGATGAGTCTGACGTGTGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3903
Db 13334 AGAGATGAGGCTCACCATGTTGCTCAGAGTGTGCTGAGACTCTGAGCTCAAGTATCC 13393
QY 3904 CCCTGCTTACCTCCCAAGGAGGCTGAGTATACAGTGTGAGTGTGAGTGTGAGTGTGAG 3963
Db 13394 ATCTTCTGCGCTGCCCAAGTGTGAGTATAGGATTAAGCACTGACCTGAGTCTCC 13453
QY 3964 CACAT-----GGTACTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGGGTTTCACT 4012

Db 3398 AACCTAAATATGTGT-----GGGCCAGGCGCAGTGGCTCAGCGCTTAATCCCAACACTT 3344
QY 4378 TGAGGGGCGCAAGGTGGGGGATCATCTTGAAGCCCTGGAGTT-----CAGCGTGGGCAACA 4431
Db 3343 TGGGAGGCGCGAGGCGGCGGATCATCTGAGGTGAGGATTCAGACAGCGCTGGCCACA 3284
QY 4432 TAGTGAGACCCGCTCTCTACCAAAATTTAAAAATTTAGCTGGAGTGGTGGCATTTGCC 4491
Db 3283 CAGTAAACCCGCTCTCTA-CAAAAATAACAAATACTGGGCAATGATGGCAGGTGCC 3225
QY 4492 TGTGTCCAGCTACTTGGGAGAGTGGGATGGGGATGGCTGAAGCTGTGAGTGGAG 4551
Db 3224 TGTAAATCCAGCTACTTGGGAGAGTGGGATGGGATGGCTGAATCTGGGAGGCGGAG 3165
QY 4552 GCTGAGAGTGAATGATGATCAGACCACTGCACTTCAAGCTGATGAGTGAAGCTATCTCAAA 4611
Db 3164 GTTGAGAGTGGCGGAGATGATGAGCGGCTGCACTGCAACCTGGGAGACAGAGAACTAGG 3105
QY 4612 GCAACAAATAATATGTTTATCTTAA 4636
Db 3104 TCTCTCTCAAAAGATTTAAAAAAA 3080

RESULT 13

US-09-764-846-335/c
; Sequence 335, Application US/09764846
; Patent No. US20020102638A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic acids, proteins, and antibodies
; FILE REFERENCE: PT12
; CURRENT APPLICATION NUMBER: US/09/764,846
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 335
; LENGTH: 10139
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-846-335

Query Match 6.6%; Score 329; DB 10; Length 10139;
Best Local Similarity 69.6%; Pred. No. 1.3e-80;

Matches 560; Conservative 0; Mismatches 210; Indels 35; Gaps 7;

QY 3838 AGAGAGAGATGAGATGCTCACTGTGTGTCAGAGCGTGTGCAACTCTAGGCGCAAG 3897
Db 3855 AGAGACAGAGTGTGTCATGTCATGATGGTGTGATGGCTGTTCGAATCTTGGCTCAAG 3796
QY 3898 CAATCCCTGCTTACGCTCCCAAGGGCTGGGATTAAGGTGATGAGTACTGCACTTG 3957
Db 3795 TGATCTGCCACACCGACCTACCAAACTGTAAATTAAGCATGAGCATGAGGCACTG---TG 3740
QY 3958 ACCAAGCAGATGATGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4017
Db 3739 TCCAGTCTCATATATATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3680
QY 4018 ACCAGGCTGAGTGAAGTGGGGCAATCTGCTCACTGTAACCTGAGCTGCCGCCAGGTG 4077
Db 3679 TGCCAGGTTGAGATGCAATGACACAGTCTTGCTCACTGCAACCCCTGCTCCAGGTT 3620
QY 4078 CAGCGATTCCTGCTTACGCTCCCTGAGTGAATTAAGACACAGGTTTCAATCATTTGGC 4137
Db 3619 CAAACATTCCTGCTCAGCTCCACCTCCCAAGTTGCGGAATTAACAGCCGCCGCCAGACA 3560
QY 4138 CCGGCTAATTTTTTTTTTTTGT 4197
Db 3559 CCGGCTAAT-----TTTTGTGAATTTTAAAGAGACAGATGATGGCATTTGGC 3509
QY 4198 CAGCGTGTCTGAACCCCTGACCTCAAGTATGATCCAGCCAGCTGGGCTCCCAAGTGTCT 4257
Db 3508 CAGCGTGTCTCAAACTCTCACTCAGGCGATCCAGCCGCTTGGCT-CCAAAGTGTCT 3450

QY 4258 GGGATTACAGGTGTGACGCCACATGACAGCCACATGATGATTTTAAATTTATTTT 4317
Db 3449 GGGATTACAGGTGTGACGCCAC-----CGACCTGGCCATTTACGTTAATTTTAT 3399
QY 4318 TTAATTAATGTTTATCTAAGCCAGTACAGTACTGCGTGTGTAATCCAGCACTT 4377
Db 3398 AACCTAAATATGTGT-----GGGCCAGGCGCAGATGGCTCAGCGCTGAATCCCAACTT 3344
QY 4378 TGAGGGGCGCAAGTGGGGGATCATCTTGAACCTGGAGTT-----CAGCGTGGGCAACA 4431
Db 3343 TGGGAGGCGGAGGCGGAGATCATCTGAGGTGAGGATTCAGACAGCGCTGGCCACA 3284
QY 4432 TAGTGAGACCCGCTCTCTACCAAAATTTAAAAATTTAGCTGGGAGTGGGATTTGCC 4491
Db 3283 CAGTAAACCCGCTCTCTA-CAAAAATAACAAATACTGGGCAATGATGGCAGGTGCC 3225
QY 4492 TGTGTCCAGCTACTTGGGAGAGTGGGATGGGGATGGCTGAAGCTGTGAGGTGAG 4551
Db 3224 TGTAAATCCAGCTACTTGGGAGAGTGGGATGGGATGGCTGAATCTGGGAGGCGGAG 3165
QY 4552 GCTGAGAGTGAATGATGATCAGACCACTGCACTTCAAGCTGATGAGTGAAGCTATCTCAAA 4611
Db 3164 GTTGAGAGTGGCGGAGATGATGAGCGGCTGCACTGCAACCTGGGAGACAGAGAACTAGG 3105
QY 4612 GCAACAAATAATATGTTTATCTTAA 4636
Db 3104 TCTCTCTCAAAAGATTTAAAAAAA 3080

RESULT 14

US-09-860-670-250/c
; Sequence 250, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic acids, proteins, and antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 250
; LENGTH: 14448
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-860-670-250

Query Match 6.6%; Score 328; DB 9; Length 14448;
Best Local Similarity 69.8%; Pred. No. 3.2e-80;

Matches 561; Conservative 0; Mismatches 175; Indels 68; Gaps 6;

QY 3844 AGAGATGAGATGCTCACTGTGTGTCAGAGCGTGTGCAACTCTAGGCTCAAGCAATCC 3903
Db 9307 AAAGATGGGTCTCATATATGTTGCGGAGCGGCTTGAATCTGCTCAAGTGTATCC 9248
QY 3904 CCGTGCCTTACGCTCCCAAGGGGCTGGGATTAAGGTGATGAGCTGATGCACTTACCAAC 3963
Db 9247 TCTGCTCTGGGCTCTTAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9191
QY 3964 CACATGTAATTTTTTTTTTTTTTTTTTTTTTTTGAAGACAGGTTTCACTC-CATCAACCA 4022
Db 9190 -----CGAATTAATTTTAAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9136
QY 4023 GCGTGAAGTCAAGTGGGGCAATCTGGCTCACTGTAACCTGTGCTCCAGGTGCAAGC 4082
Db 9135 GCGTGAAGTCAAGT-GGCAAGATCTGCGCTGACCCCAACCTCACTCCCGGTTCAAGC 9077
QY 4083 GATTCCTGCGCTTACCTCCGAGTACGTGAATTAAGGACACACCAAGCTGTG 4142
Db 9076 GATTCCTGCGCTCAAGCTCTGAGTACGTGGGATTAAGGCAATGACACCAACCA 9017


```

QY 4143 CTAATTTTTTTTTTTCTGATATTTTATAGAGACAGGGTTTCATCATGTTGGCCAGGC 4202
    |||||
Db 9016 ATAAATT-----TCATATTTTAGTAGAGACAGGGTTTCTCCACATTTGGTCAAGC 8968
    |||||
QY 4203 TGTCTTTGAACCCCTGACCTCAAGTATCCACACCTCGGCTCCCAAAAGTGTGGAT 4262
    |||||
Db 8967 TGTCTTTGAACCCCTGACCTCAAGTATCCACACCTCGGCTCCCAAAAGTGTGGAT 8908
    |||||
QY 4263 TACAGGTGTGACGACCATGACACGACATGTTACATTTTAAATTTATTTTAA- 4321
    |||||
Db 8907 TACAGGATGTGACGACCATGACACGACATGTTTCTAAATTTTAAATTTTAAAGAG 8848
    |||||
QY 4322 -----TTAAATGTTATCTAAAGC 4341
    |||||
Db 8847 GACCCAGAGATCAAAAGTGCCTATTAACCCAGAAAGTCTAAATGAGACAAAGGAGC 8788
    |||||
QY 4342 CAGTAGCAGTACCTCGGCTGTATATCCACACCTTTGAGGGCCAAAGTGGGGATCA 4401
    |||||
Db 8787 TGGGCGTGTGCTCACTCTGTATATCCACACCTTTGAGGGCCAAAGTGGGGATCA 8728
    |||||
QY 4402 CTGAGCCTGGAGTT-----CAGCGTGGGCAACATAGTGAACCCCGCTCTACGAA 4455
    |||||
Db 8727 CCGAGGTGAGAGTGTGAGACACGCTGGCCACATATGTAACCTGTCTCTACTAA 8668
    |||||
QY 4456 AATTTAAAAAATTAGCTGGAGTGTGCGCATTTGCTGTGCTCCACACTACTTGGGAGC 4515
    |||||
Db 8667 AAACACAAAAAATTAGTGTGCTGTGCTGTGCTGGGCGACCATATGATCACTCTCTCAAGAGG 8608
    |||||
QY 4516 TGAAGTGTGGGATGCTGAAGCTGTGAGCTGAGAGCTCACTGAGCTATGATCACACC 4575
    |||||
Db 8607 TGAAGGAGAGAAATCGCTTGAACCTAGAGGACAGAGCTCAGTGGGCTGAGATCGACCC 8548
    |||||
QY 4576 ACTGCACTTCAGCCTGAGTGCAG 4599
    |||||
Db 8547 ACTGCACTTCAGCCTGAGTGCAG 8524
    |||||

```

RESULT 15
 US-09-860-253/c
 ; Sequence 253, Application us/09860670
 ; Patent No. US20020165137A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PA127P1
 ; CURRENT APPLICATION NUMBER: us/09/860,670
 ; CURRENT FILING DATE: 2001-05-21
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 289
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 253
 ; LENGTH: 14451
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-860-253

Query Match 6.6%; Score 328; DB 9; Length 14451;
 Best Local Similarity 69.8%; Pred. No. 3.2e-80;
 Matches 561; Conservative 0; Mismatches 175; Indels 68; Gaps 6;

```

QY 3844 AAGATGAGAGTCTCACTGTGTTGTCAGGCTGCTCGAATCTTAGGCTCAAGCAATCC 3903
    |||||
Db 9307 AAGATGAGAGTCTCACTGTGTTGTCAGGCTGCTCGAATCTTAGGCTCAAGCAATCC 9248
    |||||
QY 3904 CCTGCTTAGGCTCCAGAGGGGCTGGATTAACAGTGTGAGCTACTGCACTTGACCAAC 3963
    |||||
Db 9247 TCTGCTTAGGCTCCAGAGGGGCTGGATTAACAGTGTGAGCTACTGCACTTGACCAAC 9191
    |||||
QY 3964 CACATGCTACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4022
    |||||
Db 9190 -----CGAATATATATTTTAAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9136
    |||||
QY 4023 GGCTGAGTGCAGTGGGGCAATCTTGCTCACTGTAACCTGCTGCCCTCCAGGTGCAAGC 4082
    |||||

```

```

Db 9135 GGCTGAGTGCAGTGGGGCAATCTTGCTCACTGTAACCTGCTGCCCTCCAGGTGCAAGC 9077
    |||||
QY 4083 GATTCCTCGCTTAGGCTCCTCTGATGATGCTGAAATTAATAGGCACACACACACGCTGG 4142
    |||||
Db 9076 GATTCCTCGCTTAGGCTCCTCTGATGATGCTGAAATTAATAGGCACACACACACGCTGG 9017
    |||||
QY 4143 CTAATTTTTTTTTTTCTGATATTTTATAGAGACAGGGTTTCATCATGTTGGCCAGGC 4202
    |||||
Db 9016 ATAAATT-----TCATATTTTAGTAGAGACAGGGTTTCTCCACATTTGGTCAAGC 8968
    |||||
QY 4203 TGTCTTTGAACCCCTGACCTCAAGTATCCACACCTCGGCTCCCAAAAGTGTGGAT 4262
    |||||
Db 8967 TGTCTTTGAACCCCTGACCTCAAGTATCCACACCTCGGCTCCCAAAAGTGTGGAT 8908
    |||||
QY 4263 TACAGGTGTGACGACCATGACACGACATGTTTAAATTTTAAATTTTAA- 4321
    |||||
Db 8907 TACAGGATGTGACGACCATGACACGACATGTTTCTAAATTTTAAATTTTAAAGAG 8848
    |||||
QY 4322 -----TTAAATGTTATCTAAAGC 4341
    |||||
Db 8847 GACCCAGAGATCAAAAGTGCCTATTAACCCAGAAAGTCTAAATGAGACAAAGGAGC 8788
    |||||
QY 4342 CAGTAGCAGTACCTCGGCTGTATATCCACACCTTTGAGGGCCAAAGTGGGGATCA 4401
    |||||
Db 8787 TGGGCGTGTGCTCACTCTGTATATCCACACCTTTGAGGGCCAAAGTGGGGATCA 8728
    |||||
QY 4402 CTGAGCCTGGAGTT-----CAGCGTGGGCAACATAGTGAACCCCGCTCTACGAA 4455
    |||||
Db 8727 CCGAGGTGAGAGTGTGAGACACGCTGGCCACATATGTAACCTGTCTCTACTAA 8668
    |||||
QY 4456 AATTTAAAAAATTAGCTGGAGTGTGCGCATTTGCTGTGCTCCACACTACTTGGGAGC 4515
    |||||
Db 8667 AAACACAAAAAATTAGTGTGCTGTGCTGTGCTGGGCGACCATATGATCACTCTCTCAAGAGG 8608
    |||||
QY 4516 TGAAGTGTGGGATGCTGAAGCTGTGAGCTGAGAGCTGCACTGAGCTATGATCACACC 4575
    |||||
Db 8607 TGAAGGAGAGAAATCGCTTGAACCTAGAGGACAGAGCTCAGTGGGCTGAGATCGACCC 8548
    |||||
QY 4576 ACTGCACTTCAGCCTGAGTGCAG 4599
    |||||
Db 8547 ACTGCACTTCAGCCTGAGTGCAG 8524
    |||||

```

Search completed: June 9, 2003, 05:30:55
 Job time : 678.38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 10:29:42 ; Search time 206.047 Seconds
(without alignments)
7441.920 Million cell updates/sec

Title: US-09-622-964-1_COPY_1_5000

Perfect score: 5000

Sequence: 1 ccaaaaattgtctcttgg.....tttggaagccgagcgaggt 5000

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :

1: Issued Patents NA: *
2: /cgn2_6/prodata/1/lna/5A_COMB.seq: *
3: /cgn2_6/prodata/1/lna/5B_COMB.seq: *
4: /cgn2_6/prodata/1/lna/5C_COMB.seq: *
5: /cgn2_6/prodata/1/lna/5D_COMB.seq: *
6: /cgn2_6/prodata/1/lna/5E_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------------|-------------------|
| 1 | 334.6 | 6.7 | 14581 | 4 US-08-520-373D-4 | Sequence 4, Appl |
| 2 | 334.6 | 6.7 | 22481 | 4 US-08-367-841A-43 | Sequence 43, Appl |
| 3 | 334.6 | 6.7 | 22481 | 5 PCT-US95-07201-43 | Sequence 43, Appl |
| 4 | 334.6 | 6.7 | 22484 | 4 US-09-875-223-2 | Sequence 2, Appl |
| 5 | 327.4 | 6.5 | 246240 | 2 US-08-724-394A-20 | Sequence 20, Appl |
| 6 | 327.4 | 6.5 | 246240 | 2 US-08-724-394A-21 | Sequence 21, Appl |
| 7 | 327.4 | 6.5 | 246240 | 2 US-08-724-394A-22 | Sequence 22, Appl |
| 8 | 303.2 | 6.1 | 43950 | 4 US-09-735-934A-3 | Sequence 3, Appl |
| 9 | 301.4 | 6.0 | 7676 | 1 US-08-451-777A-7 | Sequence 7, Appl |
| 10 | 301.4 | 6.0 | 7676 | 2 US-08-451-777A-7 | Sequence 7, Appl |
| 11 | 301.4 | 6.0 | 7676 | 2 US-08-998-208-7 | Sequence 7, Appl |
| 12 | 301.4 | 6.0 | 7676 | 5 PCT-US95-06743-7 | Sequence 7, Appl |
| 13 | 299.8 | 6.0 | 246240 | 2 US-08-724-394A-20 | Sequence 20, Appl |
| 14 | 299.8 | 6.0 | 246240 | 2 US-08-724-394A-21 | Sequence 21, Appl |
| 15 | 299.8 | 6.0 | 246240 | 2 US-08-724-394A-22 | Sequence 22, Appl |
| 16 | 298.4 | 6.0 | 99300 | 4 US-09-798-096-10 | Sequence 10, Appl |
| 17 | 296.8 | 5.9 | 43950 | 4 US-09-735-934A-3 | Sequence 3, Appl |
| 18 | 293.8 | 5.9 | 84495 | 4 US-09-797-906-3 | Sequence 3, Appl |
| 19 | 289.6 | 5.8 | 98844 | 4 US-09-791-211-10 | Sequence 10, Appl |
| 20 | 289.6 | 5.7 | 59065 | 4 US-09-813-817-3 | Sequence 3, Appl |
| 21 | 285.2 | 5.7 | 59065 | 4 US-09-878-197-3 | Sequence 3, Appl |
| 22 | 282.8 | 5.7 | 14581 | 4 US-08-520-373D-4 | Sequence 4, Appl |
| 23 | 282.8 | 5.7 | 22481 | 4 US-08-367-841A-43 | Sequence 43, Appl |
| 24 | 282.8 | 5.7 | 22481 | 5 PCT-US95-07201-43 | Sequence 43, Appl |
| 25 | 282.8 | 5.7 | 22484 | 4 US-09-875-223-2 | Sequence 2, Appl |
| 26 | 279.8 | 5.6 | 98844 | 4 US-09-791-211-10 | Sequence 10, Appl |
| 27 | 277.2 | 5.5 | 12394 | 4 US-09-488-856A-10 | Sequence 10, Appl |

| | | | | | |
|------|-------|-----|--------|---------------------|--------------------|
| C 28 | 275 | 5.5 | 40000 | 4 US-09-780-049-18 | Sequence 18, Appl |
| C 29 | 274 | 5.5 | 99500 | 4 US-09-798-096-10 | Sequence 10, Appl |
| C 30 | 271.6 | 5.4 | 17327 | 1 US-07-906-871-15 | Sequence 15, Appl |
| C 31 | 271.4 | 5.4 | 7680 | 4 US-09-210-748A-3 | Sequence 3, Appl |
| C 32 | 271.4 | 5.4 | 162450 | 4 US-09-345-882-1 | Sequence 1, Appl |
| C 33 | 271 | 5.4 | 17327 | 1 US-07-906-871-15 | Sequence 15, Appl |
| C 34 | 268.6 | 5.4 | 162450 | 4 US-09-345-882-1 | Sequence 1, Appl |
| C 35 | 266 | 5.3 | 7505 | 4 US-09-078-294-13 | Sequence 13, Appl |
| C 36 | 264 | 5.3 | 2561 | 4 US-09-270-542-101 | Sequence 101, Appl |
| C 37 | 264 | 5.3 | 2561 | 4 US-09-270-542-119 | Sequence 119, Appl |
| C 38 | 262.2 | 5.2 | 72604 | 4 US-09-268-992-7 | Sequence 7, Appl |
| C 39 | 262.2 | 5.2 | 72604 | 4 US-09-268-992-7 | Sequence 7, Appl |
| C 40 | 262.6 | 5.2 | 15602 | 4 US-09-844-634-17 | Sequence 17, Appl |
| C 41 | 258.8 | 5.2 | 62804 | 4 US-09-800-960-3 | Sequence 3, Appl |
| C 42 | 258.6 | 5.2 | 70000 | 4 US-09-851-896-3 | Sequence 3, Appl |
| C 43 | 257.4 | 5.1 | 4803 | 4 US-09-197-636-1 | Sequence 1, Appl |
| C 44 | 257.4 | 5.1 | 4803 | 4 US-09-197-636-3 | Sequence 3, Appl |
| C 45 | 256.4 | 5.1 | 15602 | 4 US-09-844-634-17 | Sequence 17, Appl |

ALIGNMENTS

```
RESULT 1
US-08-520-373D-4
Sequence 4, Application US/08520373D
Patent No. 6451763
GENERAL INFORMATION:
APPLICANT: Tombran-Tink, Joyce
APPLICANT: Steele, Plunk R
APPLICANT: Chader, Gerald J
APPLICANT: Becerra, Sofia P
APPLICANT: Johnson, Lincoln V
APPLICANT: Rodriguez, Ignacio R
TITLE OF INVENTION: RETINAL PIGMENTED EPITHELIUM DERIVED NEUROTROPIC FACTOR
FILE REFERENCE: 2026-4203US1
CURRENT APPLICATION NUMBER: US/08/520,373D
CURRENT FILING DATE: 1995-08-29
PRIOR APPLICATION NUMBER: 08/377,710
PRIOR FILING DATE: 1995-01-25
PRIOR APPLICATION NUMBER: 08/279,979
PRIOR FILING DATE: 1994-07-25
PRIOR APPLICATION NUMBER: 07/894,215
PRIOR FILING DATE: 1992-06-04
PRIOR APPLICATION NUMBER: 07/952,796
PRIOR FILING DATE: 1992-09-24
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 14581
TYPE: DNA
ORGANISM: HUMAN
FEATURE:
OTHER INFORMATION: mRNA: 6683; EXON: 6683-6790; EXON 11584-11675;
OTHER INFORMATION: EXON: 14539-14581; INTRON: 6791-11583; INTRON:
OTHER INFORMATION: 11676-14538; CDS: 11584-11675; 14539-14580
US-08-520-373D-4
Query Match 6.7% Score 334.6; DB 4; Length 14581;
Best Local Similarity 71.2%; Pred. No. 1.2e+80;
Matches 549; Conservative 0; Mismatches 169; Indels 53; Gaps 6;
DB 3844 AGAGATGAGAGTCTGCTGTGTCTGCAAGGCTGCTGCAAGTCTGAGCTCAAGCAATCC 3903
13345 AGAGATGAGAGTCTGCTGTGTCTGCAAGGCTGCTGCAAGTCTGAGCTCAAGCAATCC 13404
DB 3904 CCGTCCCTTACCTCCCAAGGCTGAGATTAAGAGTGTGAGCTACTGACCTTACCAAC 3963
13405 ATCTTCTCCGCTCCCAAGTGTGAGATTAAGAGTGTGAGCTACTGACCTTACCAAC 13464
DB 3964 CACAT-----GGTACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 4012
13465 AATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 13524
```

| | | | |
|--|-------|---|-------|
| OY | 4013 | CCATCACCGGCGGAGTGCAGTGGGGGGAACCTTGGCCATCTGTAACCTCTGCTCCC | 4072 |
| Db | 13535 | CTGTCCACCGAGCGTGGAGTACAGT--GGCACTATCTCAGCTCAGTGGAACTCTGCTCCT | 13583 |
| OY | 4073 | AGGTGCAGGCAATCTCTCCTCCTTAGCGCTCTGAGTAGTGCAGTAATATAGACACACCA | 4132 |
| Db | 13584 | GGGTTCAAAGGAATCTGTGCTCTCAGCTCTCTAGTAGTCTGGGATTTACAGCGATGCACCA | 13643 |
| OY | 4133 | CCACGCTGGCTAATTTTTTTTTTTTTTCTGTATTTTTTAGTAGACAGGGTTTCATCATG | 4192 |
| Db | 13644 | CCATGCCCCGCTAATTTTTT-----TGTATTTTAGTAGAGAC--GGTTTCACCGG | 13694 |
| OY | 4193 | TTGGCCAGGGCGGCTTGAACCCCTGACCTCAAGTATCCACCCACTCGGCTCCCAAA | 4252 |
| Db | 13695 | TTGCCCAAGATGCTCTGAACCTCTGACTCAAGTATTCACCCACTCAGCTCCCAAA | 13754 |
| OY | 4253 | GTGCTGGGATTTACAGTGTGCAGCCACCATGCACAGCCACATGTCATTTTTTAATTT | 4312 |
| Db | 13755 | GTGCTGGGATTTATAGTGTAGCGACACCTGCTA----- | 13788 |
| OY | 4313 | ATTTTAAATTTAAATTTTATCTAAGGCCATGACAGTACTCGCTCTGTATCCAG | 4372 |
| Db | 13789 | --TGTATTTTAAAAAAGTGGGTCTATGGGGCTGGGGGTGGTGCATGCTGTATTCAG | 13846 |
| OY | 4373 | CACATTAAGGGGCGCAAGTGGGGGATCACTTAGGCTG---GGAATTCAGGTGGGCA | 4428 |
| Db | 13847 | CACATTTGTGACACCGAGCGGGTGGATCACAAGTCAAGGATCGAGACATCTGCTTA | 13906 |
| OY | 4429 | ACATAGTGAGACCCCGCTCTCTACCAAAATTTAAAAATAGCTGGAGATGGTGCAATTT | 4488 |
| Db | 13907 | ACAGGGTGAACCCCGCTCTACTATAAAATACAAAATTAACCAAGCATGTGTGGGGC | 13966 |
| OY | 4489 | GCCGTGTGTCCTCCAGCTACTTGGGAAGCTAGGTGGCGATGGCTGAAGCTGTGAGCTC | 4548 |
| Db | 13967 | GCCGTGTGTCCTCCAGCTACTGCGGAGGCTGAGCGCAGGAATGGCTGGAACCTGGGAGGCG | 14026 |
| OY | 4549 | GAGGCTCAGTGAAGCTATGATCACACACACTGACCTCAGCGTGAAGAGAG | 4599 |
| Db | 14027 | GAGCTTCAATGAGCGGAGATCACGCCACCGTACTCCAGCTGAGGAGCAG | 14077 |
| RESULT 2 | | | |
| US-08-367-841A-43 | | | |
| Sequence 43. Application US/08367841A | | | |
| Patent No. 6319687 | | | |
| GENERAL INFORMATION: | | | |
| APPLICANT: Chader, Gerald J.; Rodriguez, | | | |
| APPLICANT: Ignacio R.; Mazuruk, Krzysztof; | | | |
| APPLICANT: Tombran-Tink, Joyce | | | |
| TITLE OF INVENTION: PIGMENT EPITHELIUM | | | |
| TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC | | | |
| NUMBER OF SEQUENCES: 43 | | | |
| CORRESPONDENCE ADDRESS: | | | |
| ADDRESSEE: Morgan & Flinnegan | | | |
| STREET: 345 Park Avenue | | | |
| CITY: New York | | | |
| STATE: New York | | | |
| COUNTRY: USA | | | |
| ZIP: 10154 | | | |
| COMPUTER READABLE FORM: | | | |
| MEDIUM TYPE: Floppy disk | | | |
| COMPUTER: IBM PC Compatible | | | |
| OPERATING SYSTEM: PC-DOS/MS-DOS | | | |
| SOFTWARE: WORDPERFECT 5.1 | | | |
| CURRENT APPLICATION DATA: | | | |
| APPLICATION NUMBER: US/08/367,841A | | | |
| FILING DATE: 30-DEC-1994 | | | |
| CLASSIFICATION: 435 | | | |
| PRIOR APPLICATION DATA: | | | |
| APPLICATION NUMBER: 08/257,963 | | | |
| FILING DATE: 07-JUN-1994 | | | |

[illegible]

```

OY      4489  GCGCTGTGCTCCACTACTGTGGAGAGCTGAGGATGGGTGGAAGCCTGTGAGTCC 4548
Db      13956  GCGCTGTGCTCCACTACTCGGGAGGCGCTGAGGCGAGAGATGGCTGTGAACCTTGGGAGGCG 1401
OY      4549  GAGCGTCGACGTGAGCTATGATCACACCACTGCACTTCAAGCCTGTGATGACAG 4599
Db      14016  GAGCTTGACGTGAGCGCGAGATCACGCCACCGTACTCCAGCCTGAGGCACAG 14066

RESULT 3
PCT-US95-07201-43
: Sequence 43, Application PC/TU959507201
: GENERAL INFORMATION:
:   APPLICANT: Chader, Gerald J.; Becerra, Sofia
:   APPLICANT: Patricia; Schwartz, Joan P.;
:   APPLICANT: Tanilwaki, Takayuki
:   TITLE OF INVENTION: PIGMENT EPITHELIUM
:   TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
:   NUMBER OF SEQUENCES: 43
:   CORRESPONDENCE ADDRESS:
:   ADDRESSEE: Morgan & Finnegan, L.L.P.
:   STREET: 345 Park Avenue
:   CITY: New York
:   STATE: New York
:   COUNTRY: USA
:   ZIP: 10154
:   COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy Disk
:   COMPUTER: IBM PC Compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: WORDPERFECT 5.1
:   CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: PCT/US95/07201
:   FILING DATE: 06-JUN-1995
:   CLASSIFICATION:
:   PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: 08/367,841
:   FILING DATE: 30-DEC-1994
:   PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: 08/257,963
:   FILING DATE: 07-JUN-1994
:   PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: 07/952,796
:   FILING DATE: 24-SEP-1992
:   ATTORNEY/AGENT INFORMATION:
:   NAME: DOROTHY R. AUTH
:   REGISTRATION NUMBER: 36434
:   REFERENCE/DOCKET NUMBER: 20264126PCT
:   TELECOMMUNICATION INFORMATION:
:   TELEPHONE: (212) 758-4800
:   TELEFAX: (212) 751-6849
:   INFORMATION FOR SEQ ID NO: 43:
:   SEQUENCE CHARACTERISTICS:
:   LENGTH: 22481 base Pairs
:   TYPE: Nucleic Acid
:   STRANDEDNESS: Double
:   TOPOLOGY: Unknown
:   MOLECULE TYPE: Genomic DNA
:   FEATURE:
:   NAME/KEY: Pl-147
:   LOCATION:
:   IDENTIFICATION METHOD:
:   OTHER INFORMATION: full length genomic
:   OTHER INFORMATION: sequence for PBDP plus flanking sequences.
PCT-US95-07201-43

Query Match      6.7%; Score 334.6; DB 5; Length 22481;
Best Local Similarity 71.2%; Pred. No. 1.6e-80;
Matches 549; Conservative 0; Mismatches 169; Indels 53; Gaps 6;

OY      3844  AGAAGTGAAGTCTACTGTGTGTGTCAGAGCTGTGCTCGAACTCTTAGGCTCAAGCAATCC 3903

```

| | | | |
|--|-------|--|--------|
| Db | 13334 | AAGAGATGGGGGCCCTCACCAATGTTGCTGTACAGTTGGCTGTAAATCCTGTAGACTCAAGTGAATCC | 133939 |
| QY | 3904 | CCCTGCGCTTAAGCTCTCCCAAGGGGCTGGGATTTACAGTGTGAGCTACTGCACTTGACCAAC | 3963 |
| Db | 13394 | ATCTTCCTCGGCTCGCCAAAGTGTGCGGATTTATAGGCATTAAGCCACTGACCTAGCTGCC | 134535 |
| QY | 3964 | CACAT-----GGTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGCAGGGTTTCACT | 4012 |
| Db | 13454 | AATTTTTAATATTAATATTAATTAATTTAATTAATTAATTAATTAATTTTGGAGCAGGGTCTCACT | 13513 |
| QY | 4013 | CCATCAACCCAGGCTGGAGTGCAGTGGGGGAATCTTGCGTCACTGTAACCTGTGCTCCC | 4072 |
| Db | 13514 | CTGTACCCAGGCTGGAGTGCAGT--GGACATATATCAGCTCACTGCAACCTCTGCTCCT | 135729 |
| QY | 4073 | AGGTGCACAGCATTCCTCTGCTTACCTGAGCCTCTGAGTGCAGTGAATTAATAGGCACACCA | 4132 |
| Db | 13573 | GGGTTTCAACGGAATCTGTGCTCCTGAGCCTCTGAGTGCAGTGAATTAACAGCATGCACCA | 136322 |
| QY | 4133 | CCACGCGCTGGCTAATTTTTTTTTTTTTTCTGTATTTTAAAGTAAAGAGGGTTTCAATG | 4192 |
| Db | 13633 | CCATGCGCGCTTAATTTTTT-----TGTATTTTAAAGTAAAG--GGGTTTCAACCTG | 13683 |
| QY | 4193 | TTGGCCAGGCTGGTCTTGAACCCCTGACCTCAAGTATCCACCCACCTCGGCGCTCCCAA | 4252 |
| Db | 13684 | TTGCCAGGATGTGTCTCGAATCTCTGACCTCAAGTATCCACCCACCTCAAGCTCCCAA | 137433 |
| QY | 4253 | GTCGTGGATTCAGAGTGTACAGCCACCATGCACAGCCACATGGTACATTTTTTAAAAAT | 4312 |
| Db | 13744 | GGCTGGGATTTATAGTGTGAGCCACATCGGCTGA-----137777 | |
| QY | 4313 | AATTTTTAATTAATGTTTATATCATAAGGCCATGACAGTCACTGCGTCTGTAATCCAG | 4372 |
| Db | 13778 | --TGGTTTTTAAAAAATGGGTCAATGGGGCTGGGGCGGTGGCTCATGCTGTAAATCCAG | 138355 |
| QY | 4373 | CAGTTTGAGGGGCGCAAGTGCGGGGAATCATTTGAGCGCTG-----GGAGTTCAAGCGTGGCA | 4428 |
| Db | 13836 | CAGTTTGGTAGACCGGAGCGGGTGGATCAACAAGGTCAAGGAGATGAGAACCACTCTGCTTA | 138955 |
| QY | 4429 | ACATAGTGAACCCCTCTCTTACCAAAAAATTTAAAAAATTAAGCTGGGAGTGGTGCAATTT | 4488 |
| Db | 13896 | ACACGGTGAACCCCGCTCTCTACTATAAAAAATTAACCAAGCATGTGGTGGGC | 139555 |
| QY | 4489 | GCGTGTGGTCCCAAGCTACTTGGGAGCTGAGGTGTGGGATGCTGAAAGCTGTGAGGTC | 4548 |
| Db | 13956 | GCGTGTAGTCCCAAGCTACTCGGAGGCTGAGGAGGAGAAATGCGGTGAACCTGGGAGGCG | 140155 |
| QY | 4549 | GAGGCGCAGTGAAGTATGATCAACACATGCAGATTAGCGTGAAGTGACAG | 4599 |
| Db | 14016 | GAGCTTGACGTGAGCGAGATCAAGCCACGTACTCAAGCTGAGCGACAG | 14066 |
| RESULT 4 | | | |
| US-09-875-223-2 | | | |
| ; Sequence 2, Application US/09875223 | | | |
| ; Patent No. 6391850 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: No. 6391850thwestern University | | | |
| ; APPLICANT: No. 63918501 Bouck | | | |
| ; APPLICANT: David Dawson | | | |
| ; APPLICANT: Paul Gillis | | | |
| ; TITLE OF INVENTION: Methods and Compositions for Inhibiting Angiogenesis | | | |
| ; FILE REFERENCE: 0290-2303 | | | |
| ; CURRENT APPLICATION NUMBER: US/09/875,223 | | | |
| ; CURRENT FILING DATE: 2001-06-06 | | | |
| ; PRIOR APPLICATION NUMBER: US 09/122,079 | | | |
| ; PRIOR FILING DATE: 1998-07-23 | | | |
| ; PRIOR APPLICATION NUMBER: PCT/US98/15228 | | | |
| ; PRIOR FILING DATE: 1998-07-23 | | | |
| ; PRIOR APPLICATION NUMBER: US 08/899,304 | | | |
| ; PRIOR FILING DATE: 1997-07-23 | | | |
| ; NUMBER OF SEQ ID NOS: 2 | | | |
| ; SOFTWARE: PatentIn Ver. 2.1 | | | |
| ; SEQ ID NO 2 | | | |


```

Db      2968  TTATCTAATAAAGTCTTTTATAGCGCGGGCGGTGGCTCAGCGCTTATCTCCAGCACTT
OY      4378  TGAGGGCCGAAGGTGCGGGGATCACTTGAGCCTG---GAGTTCAGCGTGGGCACATA 4433
Db      3028  TGGGAGGCCGAGGGCGGGCGGATCAAGAGGTGAGGAGATCGAAGACATCCCGGCTAAACG 3087
OY      4434  GTGAGACCCCGCTCTCTACCAAAAATTTAAAAAATTAGCTGGGAGTGTGGCATTTGCCTG 4493
Db      3088  GTTGAACCCCGCTCTCTACTATAAAAAATACAAAAAATTAGCGGGCGGTAGTGGCGGCGCTTG 3147
OY      4494  TGATGCCAGCTACTTGGGAGCTGAGGTGTGGGATGGCTGAAGCGCTGTGAGTGTGAGGC 4553
Db      3148  TAGTCCCAAGCTACTTGGGAGCTGAGGCGTGAAGAGAGATGGCGTAACCCGGGAGCGGACT 3207
OY      4554  TGCAGTGAAGCTATGATCAACACACTGCACCTTCAGCGTGAAGTGCACAGAGCTATCTCAAAAGC 4613
Db      3208  TGCAGTGAAGCGGAGATCCCGCCACTGCACCTCAGCGCTGGGGAGACAGACAGACTCCGCT 3267
OY      4614  AAACAAATAA 4624
Db      3268  TCAAAAAAAAA 3278

RESULT 6
US-08-724-394A-21
/ Sequence 21, Application US/08724394A
/ Patent No. 5872237
/ GENERAL INFORMATION:
/ APPLICANT: Feder, John N.
/ APPLICANT: Krommal, Gregory S.
/ APPLICANT: Lauer, Peter M.
/ APPLICANT: Ruddy, David A.
/ APPLICANT: Thomas, Winston
/ APPLICANT: Tsuchihashi, Zenta
/ APPLICANT: Wolff, Roger K.
/ TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
/ NUMBER OF SEQUENCES: 31
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: TOWNSEND AND TOWNSEND AND CREM LLP
/ STREET: Two Embarcadero Center, 8th Floor
/ CITY: San Francisco
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94111-3634
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: IBM PC compatible
/ SOFTWARE: Patent Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/724,394A
/ FILING DATE: 01-OCT-1996
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Flits, Renee A.
/ REGISTRATION NUMBER: 35,136
/ REFERENCE/DOCKET NUMBER: 017957-000100
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-576-0200
/ TELEFAX: 415-576-0300
/ INFORMATION FOR SEQ ID NO: 21:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 246240 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: not relevant
/ TOPOLOGY: not relevant
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: 1..246240
/ OTHER INFORMATION: /note= "HLA-H.CONTIG"

```

| Query Match | Best Local Similarity | 72.9% | Score 327.4 | DB 2 | Length 246240 |
|-------------|-----------------------|--|----------------|-----------|---------------|
| Matches 489 | Conservative | 0 | Mismatches 151 | Indels 31 | Gaps 4 |
| OY | 3974 | TTTTTTTTTTTTTTTTTTTTTTTTTTTGTGAGACAGGGTTTCACTCCATCATCCAGGCTGGAGTGC | 4033 | | |
| DB | 2619 | TTTTATTTTATTTTATTTTATTTTGTGAGACGTTGTCTCACACCTGTGTGCCAGGCTGGAGTGC | 2678 | | |
| OY | 4034 | AGTGGGGCAATCTTGCTCACTGTAACTTGCCTGCCAGGTGCAGACGATTTCTCTGC | 4093 | | |
| DB | 2679 | AGT-GGAGCACTTGCTGTGCTCACTGCAATCCCAAGTCCAGGTTCAACAGATTCCTCTGC | 2733 | | |
| OY | 4094 | CTTAGCCCTCCAGTAGCTGGAAATTTATAGACACACACACACACGCTGGCTATTTT | 4155 | | |
| DB | 2738 | CTCAGTCTCCCAACACAGCTGGGATTTACAGGTGCTGCCACCATCCCACTAAATTT-- | 2794 | | |
| OY | 4154 | TTTTTCTGTATTTTATAGTAGACAGGGTTTCAATGTTGGCCAGGCTGTCTTGAAC | 4213 | | |
| DB | 2795 | -----TGATTTTATAGTAGACAGGGTTTCCCGTGTGGCAGGCTGTCTCGAAC | 2844 | | |
| OY | 4214 | CCCTGACCTCAAGGATTCACCCACCTTGGCTCCCAAAAGTCTGGGATTCAGAGTGTCA | 4273 | | |
| DB | 2848 | TCCGACCTCAAGGATTCACCCGCTGGCTCCCAAAAGTCTGGGATTCAGAGTGTGA | 2907 | | |
| OY | 4274 | GCCACCATGACACACCCACATGTATCAAT-----TTTTAAATTAATTTT | 4317 | | |
| DB | 2908 | GCCACTGCGCCACACCGGGGTTTCAATCTTATATATATATAGATTTATATTTCTGTT | 2967 | | |
| OY | 4318 | TTAATTTAAATGTTTATCTTAAAGCCAGTAGACAGTACTGCTGTGTATCTCCAGCACTT | 4377 | | |
| DB | 2968 | TTATCTAAAAAGCTTTTATTAAGGCCGGGGCGGTGTCTCACGCTGTATCTCCAGCACTT | 3027 | | |
| OY | 4378 | TGAGGGGCAAGGTGGCGGGATCTCACTTGAAGCTG-----GGAGTTCAGGGTGGGCAACTA | 4433 | | |
| DB | 3028 | TGGAGAGCCGAGGCGGGGCGGATCTCACAGAGTTCAGAGATTCGAGACATCCCGCTTAACG | 3087 | | |
| OY | 4434 | GTCGAGCCCGCTCTTACCAAAAATTTAAAAATTAAGCTGGAGTGTGTGCTTTGCTGTG | 4493 | | |
| DB | 3088 | GTCGAGCCCGCTCTCTACTTAAAAATTAACAAAAATTAAGCGGGGCGGTAGTGGCGGGCGCTG | 3147 | | |
| OY | 4494 | TGCTCCAGCTACTTGGGGAAGCTAAGTGTGGGATGGCTGACACCTGTGAGTCTGAGGC | 4553 | | |
| DB | 3148 | TAGTCCAGCTACTTGGGAGGCTGAGGACAGAGATGGCGTAACCCGGGAGGCGAGCT | 3207 | | |
| OY | 4554 | TGCAGTAGCTATGATCACACACACTGTGACTTACCTGAGTGAAGCAGGCTATCTCAAAAGC | 4613 | | |
| DB | 3208 | TGCAGTAGCGGAGATCTCCCGCACTGCATCTCAGAGCTTGGCGGACAGACAGCAAGTCCGCTC | 3267 | | |
| OY | 4614 | AAACAAATAA 4624 | | | |
| DB | 3268 | TCAAAAAAAAAA 3278 | | | |

RESULT 7

US-08-724-394A-22

Sequence 22, Application US/08724394A

Patent No. 5872237

GENERAL INFORMATION:

APPLICANT: Feder, John N.

APPLICANT: Krommal, Gregory S.

APPLICANT: Lauder, Peter M.

APPLICANT: Ruddy, David A.

APPLICANT: Thomas, Winston

APPLICANT: Tsuchinashi, Zenta

APPLICANT: Wolff, Roger K.

TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1

TITLE OF INVENTION: Sequences and Antibodies Thereto

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW LLP

STREET: Two Embarcadero Center, 8th Floor

```

1  CTRY:  San Francisco
2  STATE:  CA
3  COUNTRY:  USA
4  ZIP:  94111-3834
5  COMPUTER READABLE FORM:
6  MEDIUM TYPE:  Floppy disk
7  COMPUTER:  IBM PC compatible
8  OPERATING SYSTEM:  PC-DOS/MS-DOS
9  SOFTWARE:  Patentn Release #1.0, Version
10 CURRENT APPLICATION DATA:
11 APPLICATION NUMBER:  US/08/724,394A
12 FILING DATE:  01-OCT-1996
13 CLASSIFICATION:  536
14 ATTORNEY/AGENT INFORMATION:
15 NAME:  Flits, Renee A.
16 REGISTRATION NUMBER:  35,136
17 REFERENCE/DOCKET NUMBER:  017957-000100
18 TELECOMMUNICATION INFORMATION:
19 TELEPHONE:  415-576-0200
20 TELEFAX:  415-576-0300
21 INFORMATION FOR SEQ ID NO:  22:
22 SEQUENCE CHARACTERISTICS:
23 LENGTH:  246240 base pairs
24 TYPE:  nucleic acid
25 STRANDEDNESS:  not relevant
26 TOPOLOGY:  not relevant
27 MOLECULE TYPE:  CDNA
28 FEATURE:
29 NAME/KEY:  misc_feature
30 LOCATION:  1..246240
31 OTHER INFORMATION:  /note="HLA-H.CONFIG"
32 US-08-724-394A-22

```

QY 4372 GCACCTTGAAGGCGCAAGTGGCGGATCACTTGAGCTGGAGTTCAG-----CGTGG 4425
|||||
Db 5904 ACACCTTTAGAGAGACCGAGCGGCGGACATCACTGAGTCGGAGTTTGAACCATCCCTGG 5963
4426 GCACATAGTGAAGACCCGCTCTCTACCAAAATTTAAATTAAGTCGGAGTGGTGGCA 4485
|||||
Db 5964 CCACATAGTGAAGACCCGCTCTCTACCAAAATTTAAATTAAGTCGGAGTGGTGGCG 6022
4486 TTTGCTGTGCTGCTCCAGCTACTTGGAGAGTGTGGAGTGGAGTGGTGAAGCTGTGAG 4545
6023 CAGGCTGTGAATCTCTGAGCTGAGGCTGAGGAGAACTGCTTGAACCCGGGAG 6082
QY 4546 GTGAGGCTGAGTGAAGTATGATCAACACCTGACCTTGAAGCTGTGAGTGAAG 4599
6083 G-CGGACTGTGTGAGGCGGAGATCTCACTACTGACTCCAGCCTGGGTGACAG 6135
Db
RESULT 9
US-08-451-777A-7
Sequence 7, Application US/08451777A
Patent No. 5789223
GENERAL INFORMATION:
APPLICANT: Bergsma, Derk J.
APPLICANT: Stambolian, Dwight
TITLE OF INVENTION: Human Galactokinase Gene
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corp./Corporate
ADDRESSEE: Intellectual Property
STREET: 709 Swedeland Road/DW2220
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,777A
FILING DATE: 26-MAY-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10825
FILING DATE: 23-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Eagle, Alissa M.
REGISTRATION NUMBER: 37,126
REFERENCE/DOCKET NUMBER: P50268-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5364
TELEFAX: 610-270-5090
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 7676 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-451-777A-7
Query Match 6.0%; Score 301.4; DB 1; Length 7676;
Best local Similarity 52.6%; Pred. No. 9.2e-72;
Matches 1052; Conservative 0; Mismatches 841; Indels 106; Gaps 14;
QY 2712 ATTTATTTATTTATATCTTAAAGACAGAGTTTGGCTCTCTTACCCAGGCTGTGAGTGA 2771
Db 3160 ATTTGCGGTGTTTTTTTGGACACAGAGTTTGGCTCTCTTACCCAGGCTGTGAGTGA 3219
QY 2772 ATGGCGTATCTGAGCTGACAGCTCCACCTCGGAGTCAAGAGATTCCTGCGCT 2831
Db 3220 ATGGTGCATCTCAGCTCAGTCACTGACCTTCCCGGGGTCAAGTGAATTCCTGCGCT 3279

QY 2832 CAGCTCTCTGATGACTGGGATTTACAGGACCCCGACACCGCTGATTAATTTTGTGA 2891
|||||
Db 3280 CAGCTCTCCGATGAGTGGGATGAGGAGCGCGCCGACCTAGCCCTGGACATTTTGTATA 3339
2892 TTTTATAGAGAC--AGGGTTTACCATGTTGGGACGGCTGTGCAAGTCCGACCTT 2950
3340 GTTCTAGTGAAGATGGGGTTTACCATGTTGGGACGGCTGTGCAAGTCCGACCTT 3399
QY 2951 AGGTATTCACCTCTCTGACTTCCCAAGTGTGGATTTATAGCATAGGACCTAGCGC 3010
3400 AGGTATTCACCTCTCTGACTTCCCAAGTGTGGATTTATAGCATAGGACCTAGCGC 3459
QY 3011 C-----CAGTATTTAGAAAGTTAAAGGCATGTGCATGACACCGCTATGTA 3060
3460 CGGTCTCTTTTTTTTTTTTTTAACTAAACATTAATCTCAGAACCCGAAACCTATCTT 3519
Db
QY 3061 CGTCTTCCCTGCCCAAGCAAGGCGAGCTCTGGGCTCAGTCTTGTGGTTTCTACTTCC 3120
3520 ATCTTATGCAATGAAAGGCAATATCTCGGCTGTGCTCTTTTTTTTTTTTTTTTTTTT 3579
QY 3121 AAAAGCAGTCAGAACTGCGAGGCGCTTGGAGCCACTTCACTCACTCTAGGGTCCCT 3180
3580 TGGCGAGGTGGAGAGCTTGCCCTGTGGCCAGGCTGGAGTGCAGGGGCAATCTCGGTT 3639
QY 3181 ATGGAGAGTTGAGAGTCCAGACAGCAGGAGGGTCTGACAGGCTGTACACAGGCTCTG 3240
3640 CACTGCAATCTCCACCTCTGCGGTCCAAAATGATCTCTGCTTGAAGTATAGT 3699
QY 3241 ATCCCTCAAAACCCCAATGCGTGTGCTCTCAACAGGACCAAGCCACTGCTGAG 3300
3700 GGGATTTCTGGAACCCACACACAGCCAGCCAAATTTTATATTTTATGTAAGCGGG 3759
QY 3301 CCCACTGCTGCGCATGACATCACTTACACAGCAAGTGGCTAATGCCCTAGGCT 3360
3760 TTTTATG--TTGGCAGGCTGCGCTTCAAGTCTTACCTGATGCTGCGCCCTAGGCT 3818
QY 3361 CTTTCTCCCGCTCTGCTGTGCTGGCGGGGACAGCTTACAGCTG--CTATATGGCG 3417
3819 CCCAATGTGTAGATTTACATGTGTGAGCCACTGACCTGCGCTGCTGCTTTAA 3878
QY 3418 AGTTCCTAATCTCTGCTGCTGCTACATCAATCCGTTTATTTATAGTAAAGCTGGC 3477
3879 AGCTCCACAAATTTTACATGATGCTGCTGCTGCTGCTTACTTAACTATTTCTGATACA 3938
QY 3478 AGGCTGGCGGGGGGCTGGGAAAGATGAGTGGGCTGGGAGCTGGAGCTCCG 3537
3939 CTTCACTGCGATTAATCTTCTTATTTATTTTATTTATTTATTTATTTATTTATTT 3998
QY 3538 GGGCTCCCAAGCAGCTCAGGCGCCAGTGCACACAGTCCATCAACACTAAGCTGGCTC 3597
3999 TGAACAGAGTCTGCTGCTGCTGACACAGCTGAGTGAAGTGAAGTGAAGTGAAGTGA 4058
QY 3598 CTGACCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3657
4059 GCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4110
QY 3658 AGTCACTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3717
4111 AGCTAGATTTAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4170
QY 3718 CCCAGAGAGACCCCTGAGGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3777
4171 GAGGTTTACATGTTGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4230
QY 3778 GAATCCACACAGACTATAGGCCCATATGACTTATTAATAAAGAGAGAGAGAGAGAG 3837
4231 CAGGCTCCCAAGTACTGAGGATGAGCAGGATGAGCAGTGTGCTGCTGCTGCTGCTGCT 4290
QY 3838 AGAGAGAGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3897
4291 TTTTAAAGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4350

3301 CCCAGCTGCTGGCAGTACCATCATTACACAGCCAGTGGCTAATGCCCGCTAGGCT 3360
3760 TTTCATG-TTGGCCAGCTGCGCTCGAATCTCTGACCTGTGATCTGCGCCGCTACGCT 3818
3361 CTTTCTCCCGCTGCTGCTGCTGCTGGCGGAGCATCTACAGCTG---CTAATGGCG 3417
3819 CCCAATGTCTAGGATTTACATGTGTAGCAGCTGCGCTGCGCTGCTGCTTTAA 3878
3418 AGTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3477
3879 AGTCTCAATATTTTATGATTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3938
3478 AGGCGTGGCGGGGGGGGGGGGAGATGTGGCTGGGCTGGGAGCTGGGAGCTGCTG 3537
3939 CTTACACACTGCGATTTTACTTTCTTTTATTTTATTTTATTTTATTTTATTTAT 3998
3538 GGGCTCTCCAGCAGCTCAGGCGCCAGTGCACAGTCCACTACACACACTAGCTGGGCTC 3597
3999 TGAGACAGAGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4058
3598 CTGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3657
4059 GCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4110
3658 AGTACACACTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3717
4111 AGCTAGATTTAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4170
3718 CCCAGAGAGAGCCCTGAGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3777
4171 GAGGTTTATCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4230
3778 GAATCCACACAGCTCATAGGCGCCATAGTACATTTAAAGAGAGAGAGAGAGAGAGAG 3837
4231 CAGCTCCCAAGAGTGGGATGACAGGATGACAGGATGACAGGATGACAGGATGACAGGAT 4290
3838 AG 3897
4291 TTTTAAAGAGATGGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4350
3898 CAATCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3957
4351 TAAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4391
3958 ACCAACAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4017
4392 -----CACCGGTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4446
4018 ACCAGGCTGAGTGCAGTGGGGCAATCTTGGCTCACTGATGATGATGATGATGATGATGATGAT 4077
4447 GCCAGGCTGAGGCGAGT-GGCGGATCTGCGCTCACTGCAACTCCGCTTCAGGGTT 4505
4078 CAGGAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4137
4506 CAAAGCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4565
4138 COTGCTAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4197
4566 GCGAATTAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4625
4198 CAGGCT 4257
4626 CAGGCT 4685
4258 GGGATTT-----ACAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4311
4686 GGGATTTAAGGACAGAGGCTGAGCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4745
4312 TATTTT-----TAAATTAATGTTTAAAGGC 4341
4746 TCTTAGAATGATTTCTTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4805
4342 CAGTAGCAGTGCAGTGGCT-----CTGTAATCCAGCAGCTTTGAGGGCCAAAGTGGCG 4395

4806 TGGGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4865
4396 GGAATCATTGAGCTGAGGATTT-----CAGCGTGGCAACATAGTGAAGCCGCTCT 4449
4866 ACATTCATTGAGCTGAGGATTTGAGGATTTGAGGATTTGAGGATTTGAGGATTTGAGGATTTGAGGAT 4925
4450 ACCAAAAATTTAAAAATTTAGCTGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4509
4926 A--AAGAAATACAAAAATTTAGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4983
4510 GGAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4569
4984 GAGGCTGGGCTGGGAGGATTTGAGCCAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5043
4570 CACACACTGACCTGAGCTG-----AGTACAGGCTATCTCAAAAGCAACAA 4621
5044 TGCACCATGCACTGCACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 5103
4622 TAAATTTATCTAAAGCT 4640
5104 TACAGTATCTTAAAGCT 5122

RESULT 11
US-08-998-208-7
Sequence 7, Application US/08998208
Patent No. 5880105
GENERAL INFORMATION:
APPLICANT: Bergsma, Derk J.
APPLICANT: Stambolian, Dwight
TITLE OF INVENTION: Human Galactokinase Gene
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corp./Corporate
ADDRESSEE: Intellectual Property
STREET: 709 Swedeland Road/UM2220
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,208
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/451,777
FILING DATE: 26-MAY-1995
APPLICATION NUMBER: PCT/US94/10825
FILING DATE: 23-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Eagle, Ailsa M.
REGISTRATION NUMBER: 37,126
REFERENCE/DOCKET NUMBER: P50268-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5364
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 7676 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-998-208-7

Query Match 6.0%; Score 301.4; DB 2; Length 7676;
Best Local Similarity 52.6%; Fred. No. 9.2e-72;

NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corp./Corporate
ADDRESS: Intellectual Property
STREET: 709 Swedeland Road/DW2220
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06743
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10825
FILING DATE: 23-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sutton, Jeffrey A.
REGISTRATION NUMBER: 34,028
REFERENCE/DOCKET NUMBER: P50268-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5024
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 7676 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-0595-06743-7
Query Match 6.08; Score 301.4; DB 5; Length 7676;
Best Local Similarity 52.6%; Pred. No. 9.2e-72;
Matches 1052; Conservative 0; Mismatches 841; Indels 106; Gaps 14;

3181 ATGGAGAGTTGAGTCCAGACGAGGAAGGTCCTGACAGGCTGACAGGAGGCTCTG 3240
3640 CACTGATCTCCACCTCTGCTGCTCCAAATGATCTCTGCTTACCTTCTAGTAGT 3699
3241 ATCCCTAACAACCCCAATGCTGCTCTCTACAGAGCCCAAGCCCACTGCTGAG 3300
3700 GGGATTTACTGGAACCCACACACAGCCCAATTTTATTTATTTAGTAGAGAGGG 3759
3301 CCCACTGCTGGCCATGACATCACTTACAAAGCCAAAGTGGCTAATGCCGCTAGCT 3360
3760 TTTCATG-TTGGCCAGGCTGGCTCGAATCTGATCTGATCTGATCTGCTGCTGCT 3818
3361 CTTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3417
3819 CCATATGCTGATGATTAATGATGATGATGATGATGATGATGATGATGATGATG 3878
3418 AGTTCTAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3477
3879 AGCTCCAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 3938
3478 AGGCTGGGCGGGGGGCTGGGAAGATGCTGGGGCTGGGAGCTGGGAGCTCTGG 3537
3939 CCTCACACTGCAATTAATCTTCTTATTTATTTATTTATTTATTTATTTATTTAT 3998
3538 GGGCTCCAGCCCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3597
3999 TGAGACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4058
3598 CTGACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3657
4059 GCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4110
3658 AGTCACAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3717
4111 AGTAGATTAAGAGCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4170
3718 CCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3777
4171 GAGTTTATACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4230
3778 GAATCCCAACAGACTCATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3837
4231 CAGCTCCCAAAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4290
3838 AGAGAGAGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3897
4291 TTTTAAAGAGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4350
3898 CAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3957
4351 TAACTTCTGAC-TAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4391
3958 ACCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4017
4392 -----CACCGGTTATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4446
4018 ACCAGGCTGAGTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4077
4447 GCCAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4505
4078 CAAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4137
4506 CAGCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4565
4138 CTTGCTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4197
4566 GCGAGTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4625
4198 CAGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4257
4626 CAGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4685
4258 GGGATTT-----ACAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4311

```

Db      4686 GGGATTACAGGCACACGCGTGAAGCCTATACACTTCGTCTTAATAAGTT 4745
QY      4312 TATTT-----TTAATTAATGTTTATCTAAAGCC 4341
Db      4746 TCTTAGAAAATGATTCTTAGAGTAGATTACAGAGTCACAGAGCAAGATTGTAAGC 4805
QY      4342 CAGTAGAGTAGTACTCGGGT-----CTGTATCCACAGACTTTAGAGGCCAAGTGCGG 4395
Db      4806 TGSGTGGGGTGCTCACGTCTGTGCTGTATTCACACTACTTTAGAGGCCAAGTGCGG 4855
QY      4396 GGATCACTTGAGCCCTGGAGTT-----CAGCGTGGGCCACATAGTAGAGACCCTCTCT 4449
Db      4866 AGATTCAATTGAGTCAGGAATTCAGAGACCAAGCCTGGGCCAACATGGCAAACCATCTCT 4925
QY      4450 ACCAAAAATTTAAAAATTTACTGGAGTGGTGGCATTTGCCCTGTGNGCCAGACTACTG 4509
Db      4926 A--AAGAAATCAAAAATTTACCGAGGTGGTGGTGTGTGCTGTAGTCTTAGACTTA 4983
QY      4510 GGAAGCTAGGTGTGGGGATGGCTGTGAAGCTGTGAGGTGAGAGCTGACGTAGTANGAT 4569
Db      4984 GGAGGCTGGGTGGAGGATGATCAATTTAGACCCAGAGGGTTGAGACTGACGTGCTGAT 5043
QY      4570 CACACCACGTCACTTCACCTG-----AGTGCAGGCTATCTCCAACAACAAA 4621
Db      5044 TGCACCAATGAGCATCTCCACCTGGGCGCTCAAGTGTAGATCTGTCTCCAAAAA 5103
QY      4622 TAATGTTATCTTAACGCT 4640
Db      5104 TACAGTATCCTTAAGCT 5122

RESULT 13
US-08-724-394A-20/c
Sequence 20, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Kronmal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchishashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map; No. 5872237el
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs

```

| Query Match | Best Local Similarity | Score | DB 2: | Length | 246240: |
|--|-----------------------------|--------------------|--------------------|-----------------|------------|
| Matches 486; Conservative | 72.28; | Pred. No. 2.3e-70; | Mismatches 152; | Indels 35; | Gaps 6; |
| <p> TYPE: nucleic acid STRANDEDNESS: not relevant TOPOLOGY: not relevant MOLECULE TYPE: cDNA FEATURE: NAME/KEY: misc_feature LOCATION: 1..246240 OTHER INFORMATION: /note= "HLA-H.CONTIG" US-08-724-394A-20 </p> | | | | | |
| 3974 | TTTTTTTTTTTTTTTTTTTTTTTTTTT | 6.08; | DB 2: | Length | 246240: |
| 3292 | TTTTTTTTTTTTTTTTTTTTTTTTTTT | 72.28; | Pred. No. 2.3e-70; | Mismatches 152; | Indels 35; |
| 4034 | AGTGGGGGCAATCTGGCTCAGCTGTA | Score 299.8; | DB 2: | Length | 246240: |
| 3232 | AGTGGGGGCG-GATCTCGCTCAGCTG | 72.28; | Pred. No. 2.3e-70; | Mismatches 152; | Indels 35; |
| 4094 | CTTACGCTCTCGATGAGCTGGAATTA | Score 299.8; | DB 2: | Length | 246240: |
| 3173 | CTCAGCTCTCCAGTAGCTGGAGCTA | 72.28; | Pred. No. 2.3e-70; | Mismatches 152; | Indels 35; |
| 4154 | TTTTTCTGATTTTATGTAAGACAGG | Score 299.8; | DB 2: | Length | 246240: |
| 3115 | -----TGATTTTTTATGTAAGACAGG | 72.28; | Pred. No. 2.3e-70; | Mismatches 152; | Indels 35; |
| 4214 | CCCGACCTCAAGGATCCACACACAC | Score 299.8; | DB 2: | Length | 246240: |
| 3062 | TCTGACCTC-GTATCCGCTCCGCTC | 72.28; | Pred. No. 2.3e-70; | Mismatches 152; | Indels 35; |
| 4274 | GCCACCATGACACAGCCCATGTTAA | Score 299.8; | DB 2: | Length | 246240: |
| 3004 | GCCACCGGCGCCGCTTATTAAGACT | 72.28; | Pred. No. 2.3e-70; | Mismatches 152; | Indels 35; |
| 4318 | TTAATTAAATGTTTATCTAAGGCA | Score 299.8; | DB 2: | Length | 246240: |
| 2944 | TATGATTAAAGATGAACCCCGGCT | 72.28; | Pred. No. 2.3e-70; | Mismatches 152; | Indels 35; |
| 4378 | TGAGGGGCAAGTGTGGGGATCTAG | Score 299.8; | DB 2: | Length | 246240: |
| 2884 | TGGAGGCGGAGGGGGGTATCATCT | 72.28; | Pred. No. 2.3e-70; | Mismatches 152; | Indels 35; |
| 4432 | TAGTGAGACCCGCTCTCTACCAAA | Score 299.8; | DB 2: | Length | 246240: |
| 2834 | CGGCGAAACCTGCTCTCTACTTAA | 72.28; | Pred. No. 2.3e-70; | Mismatches 152; | Indels 35; |
| 4432 | TGTGTGCCACCTACTTGGGAAGCT | Score 299.8; | DB 2: | Length | 246240: |
| 2765 | TGTATATCCACGCTGTTGGAGACT | 72.28; | Pred. No. 2.3e-70; | Mismatches 152; | Indels 35; |
| 4552 | GCTGCATGAGCTATGATCACACAC | Score 299.8; | DB 2: | Length | 246240: |
| 2705 | ATTGCAGTGAAGCAAGACTGCACT | 72.28; | Pred. No. 2.3e-70; | Mismatches 152; | Indels 35; |
| 4612 | GCAACCAAAATPA 4624 | Score 299.8; | DB 2: | Length | 246240: |
| 2645 | TCTCAAAATPA 2633 | 72.28; | Pred. No. 2.3e-70; | Mismatches 152; | Indels 35; |

| | | | |
|----|------|---|------|
| OY | 4378 | TGAAGGCGCCAAAGTTCGGGGGATCACTTGACCTGGGAGTT-----CAGGTGGGCACA | 4431 |
| Db | 2884 | TGGGAGGCCCAAGGGGGTAGATCACTTGAGGTCAAGGATTGAAACCACCTGGCCAAACA | 2825 |
| OY | 4432 | TAGTGAAGACCCCGCTCTCTACCAAAAATTTAAAAATTAGCTGGAGTGTGGCATTTTCCC | 4491 |
| Db | 2824 | CGGGAAACCTGCTCTCTACTATAAAAA--TACAAAATTAAGCTGGGCATGTGTGGCAGGCCAC | 2766 |
| OY | 4492 | TGTCGTCCCAAGCTTCTTGGGAAGCTGAGGTGTGGGATGTGGCTGAAGCCTTGAGAGTCGAG | 4551 |
| Db | 2765 | TGTATTCACACTCTTGTGGAGACTGAGCGCAGAGAAATCGTTTGAACCTGGGACGTGGGG | 2708 |
| OY | 4552 | GCTCAGATGAGCTATGATCACACCACACTGCACCTTAGCCTGAGTGCAGCGCATATCTCAAAA | 4611 |
| Db | 2705 | ATTTCAGAGAACCCCAAGACTGCACACACTGCACATCCAGCCTGGGCGACAGGGGTGAGACAACG | 2644 |
| OY | 4612 | GCAACCAAAAATPAA | 4624 |
| Db | 2645 | TCTCAAAAATAA | 2633 |

```

US-08-724-394A-22/c
: Sequence 22, Application US/08724394A
: Patent No. 5872237
: GENERAL INFORMATION:
: APPLICANT: Feder, John N.
: APPLICANT: Krommal, Gregory S.
: APPLICANT: Laufer, Peter M.
: APPLICANT: Ruddy, David A.
: APPLICANT: Thomas, Winston
: APPLICANT: Tsuchihashi, Zenta
: APPLICANT: Wolff, Roger K.
: TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
: TITLE OF INVENTION: Sequences and Antibodies Thereto
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
: STREET: Two Embarcadero Center, 8th Floor
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/724,394A
: FILING DATE: 01-OCT-1996
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Fitts, Renee A.
: REGISTRATION NUMBER: 35,136
: REFERENCE/DOCKET NUMBER: 01/957-000100
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-576-0200
: TELEFAX: 415-576-0300
: INFORMATION FOR SEQ ID NO: 22:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 246240 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: not relevant
: TOPOLOGY: not relevant
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..246240
: OTHER INFORMATION: /note= "H1A-H.CONTIG"
US-08-724-394A-22

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 09:29:22 ; Search time 1005.38 Seconds
(without alignments)
11199.691 Million cell updates/sec

Title: US-09-622-964-1_COPY_1_5000
Perfect score: 5000
Sequence: 1 ccaaaattgtctctctg.....tttgagagcgagcgaggt 5000

Scoring table: IDENTITY_NDC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.GeneSeq_101002:*

| | |
|-----|--|
| 1: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1980.DAT:* |
| 2: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1981.DAT:* |
| 3: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1982.DAT:* |
| 4: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1983.DAT:* |
| 5: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1984.DAT:* |
| 6: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1985.DAT:* |
| 7: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1986.DAT:* |
| 8: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1987.DAT:* |
| 9: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1988.DAT:* |
| 10: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1989.DAT:* |
| 11: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1990.DAT:* |
| 12: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1991.DAT:* |
| 13: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1992.DAT:* |
| 14: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1993.DAT:* |
| 15: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1994.DAT:* |
| 16: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1995.DAT:* |
| 17: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1996.DAT:* |
| 18: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1997.DAT:* |
| 19: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1998.DAT:* |
| 20: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1999.DAT:* |
| 21: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2000.DAT:* |
| 22: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2001A.DAT:* |
| 23: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2001B.DAT:* |
| 24: | /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2002.DAT:* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|----------|----------------------|
| 1 | 4990 | 99.8 | 16125 | AA221226 | Human CGICe genom |
| 2 | 2758.2 | 55.2 | 16650 | ABAI4559 | Human nervous syst |
| 3 | 2758.2 | 55.2 | 18530 | ABAI4557 | Human nervous syst |
| 4 | 2758.2 | 55.2 | 18537 | ABAI4558 | Human nervous syst |
| 5 | 1291.8 | 25.8 | 6396 | AB134247 | Human immune syste |
| 6 | 1284.2 | 25.7 | 6396 | AB134246 | Human immune syste |
| 7 | 375.6 | 7.5 | 17792 | AA532727 | Human genomic DNA |
| 8 | 375.6 | 7.5 | 17792 | AA536099 | Human cardiovascular |
| 9 | 356.2 | 7.1 | 2816 | AAK83109 | Human immune/haema |

| | | | | | | |
|---|----|-------|-----|--------|----|----------|
| C | 10 | 352.6 | 7.1 | 15413 | 22 | AAK84002 |
| C | 11 | 351.4 | 7.0 | 201143 | 24 | ABK83568 |
| C | 12 | 351.2 | 7.0 | 32224 | 22 | AAK89986 |
| C | 13 | 350.6 | 7.0 | 13819 | 22 | AA136231 |
| C | 14 | 349.6 | 7.0 | 30393 | 22 | AAK67239 |
| C | 15 | 344.4 | 6.9 | 1979 | 22 | AAH17407 |
| C | 16 | 340.6 | 6.8 | 13821 | 22 | AA136230 |
| C | 17 | 338.6 | 6.8 | 9397 | 22 | AAK78065 |
| C | 18 | 337.6 | 6.8 | 149480 | 24 | AB161947 |
| C | 19 | 337.6 | 6.8 | 149480 | 24 | AB161948 |
| C | 20 | 337.6 | 6.8 | 149480 | 24 | AB161949 |
| C | 21 | 336.6 | 6.7 | 10139 | 22 | AB168365 |
| C | 22 | 335.2 | 6.7 | 19632 | 22 | AA529226 |
| C | 23 | 335 | 6.7 | 6451 | 22 | AAK66943 |
| C | 24 | 334.6 | 6.7 | 22481 | 17 | AA11658 |
| C | 25 | 334.6 | 6.7 | 22484 | 24 | AAK72461 |
| C | 26 | 334.6 | 6.7 | 22484 | 24 | AB168196 |
| C | 27 | 334.6 | 6.7 | 22484 | 24 | AB168197 |
| C | 28 | 334.6 | 6.7 | 22484 | 24 | AB168198 |
| C | 29 | 332.8 | 6.7 | 23825 | 22 | AAK79161 |
| C | 30 | 332.2 | 6.6 | 139904 | 24 | ABK83562 |
| C | 31 | 331.6 | 6.6 | 32194 | 22 | AA104340 |
| C | 32 | 330.2 | 6.6 | 11301 | 22 | AAK71591 |
| C | 33 | 330 | 6.6 | 939 | 22 | AAK83508 |
| C | 34 | 329 | 6.6 | 10139 | 22 | AA529226 |
| C | 35 | 329 | 6.6 | 175737 | 24 | ABK83571 |
| C | 36 | 328.4 | 6.6 | 220895 | 24 | ABK84798 |
| C | 37 | 328 | 6.6 | 14448 | 22 | AAK71815 |
| C | 38 | 328 | 6.6 | 14448 | 22 | AAK73111 |
| C | 39 | 328 | 6.6 | 14448 | 22 | AAK87580 |
| C | 40 | 328 | 6.6 | 14448 | 22 | AA162922 |
| C | 41 | 328 | 6.6 | 14451 | 22 | AAK71818 |
| C | 42 | 328 | 6.6 | 14451 | 22 | AAK73116 |
| C | 43 | 328 | 6.6 | 14451 | 22 | AAK87587 |
| C | 44 | 328 | 6.6 | 14451 | 22 | AA162925 |
| C | 45 | 326.4 | 6.5 | 14426 | 22 | AAK71817 |

ALIGNMENTS

| | | | | | | |
|----------|--|----------------------------|--|--|--|--|
| RESULT 1 | | | | | | |
| ID | AA221226 | standard; DNA; 16125 BP. | | | | |
| XX | XX | | | | | |
| AC | AA221226; | | | | | |
| XX | XX | | | | | |
| DT | 22-NOV-1999 | (first entry) | | | | |
| XX | XX | | | | | |
| DE | Human CGICe genomic DNA sequence. | | | | | |
| XX | XX | | | | | |
| KW | CGICe; Best's macular dystrophy; mutation; diagnosis; detection; | | | | | |
| KW | BMD; age-related macular dystrophy; ss. | | | | | |
| XX | XX | | | | | |
| OS | Homo sapiens. | | | | | |
| XX | XX | | | | | |
| FH | Key | Location/Qualifiers | | | | |
| FT | Intron | 1..1513 | | | | |
| FT | FT | /*tag= a | | | | |
| FT | FT | /*number= 1 | | | | |
| FT | FT | 1514..1570 | | | | |
| FT | FT | /*tag= b | | | | |
| FT | FT | /*number= 1 | | | | |
| FT | FT | 1571..3278 | | | | |
| FT | FT | /*tag= c | | | | |
| FT | FT | /*number= 2 | | | | |
| FT | FT | 3279..3466 | | | | |
| FT | FT | /*tag= d | | | | |
| FT | FT | /*number= 2 | | | | |
| FT | FT | 3315..15787 | | | | |
| FT | FT | /*tag= e | | | | |
| FT | FT | /*product= "CGICe protein" | | | | |
| FT | FT | /*note= "contains introns" | | | | |

| | | |
|----|---|---------------|
| FT | Intron | 3467..6616 |
| FT | | /*tag= f |
| FT | | /number= 3 |
| FT | exon | 6617..6711 |
| FT | | /*tag= g |
| FT | | /number= 3 |
| FT | Intron | 6712..7227 |
| FT | | /*tag= h |
| FT | | /number= 4 |
| FT | exon | 7228..7461 |
| FT | | /*tag= i |
| FT | | /number= 4 |
| FT | Intron | 7462..8353 |
| FT | | /*tag= j |
| FT | | /number= 5 |
| FT | exon | 8354..8508 |
| FT | | /*tag= k |
| FT | | /number= 5 |
| FT | Intron | 8509..8896 |
| FT | | /*tag= l |
| FT | | /number= 6 |
| FT | exon | 8897..8974 |
| FT | | /*tag= m |
| FT | | /number= 6 |
| FT | Intron | 8975..9676 |
| FT | | /*tag= n |
| FT | | /number= 7 |
| FT | exon | 9677..10032 |
| FT | | /*tag= o |
| FT | | /number= 7 |
| FT | Intron | 10033..11212 |
| FT | | /*tag= p |
| FT | | /number= 8 |
| FT | exon | 11213..11293 |
| FT | | /*tag= q |
| FT | | /number= 8 |
| FT | Intron | 11294..11556 |
| FT | | /*tag= r |
| FT | | /number= 9 |
| FT | exon | 11557..11708 |
| FT | | /*tag= s |
| FT | | /number= 9 |
| FT | Intron | 11709..13918 |
| FT | | /*tag= t |
| FT | | /number= 10 |
| FT | exon | 13919..14557 |
| FT | | /*tag= u |
| FT | | /number= 10 |
| FT | Intron | 14558..15768 |
| FT | | /*tag= v |
| FT | | /number= 11 |
| FT | exon | 15769..16125 |
| FT | | /*tag= w |
| FT | | /number= 11 |
| PN | WO9943695-A1. | |
| XX | | |
| PD | 02-SEP-1999. | |
| XX | | |
| XX | 22-FEB-1999; | 99WO-US03790. |
| XX | | |
| PR | 25-FEB-1998; | 98US-0075941. |
| PR | 18-DEC-1998; | 98US-0112926. |
| XX | | |
| PA | (MERI) MERCK & CO INC. | |
| PA | (UYUP-) UNIV UPSALA. | |
| XX | | |
| PI | Petrukhin K, Caskey CT, Metzker M, Madellius C; | |
| XX | | |
| XX | WPI, 1999-540560/45. | |
| DR | P-PSDB; AA129953. | |
| PT | Human and mouse polynucleotides encoding CGICE polypeptides | |

XX Claim 2; Fig 1; 67pp: English.

xx The present sequence represents the human GC1E gene, which when
xx mutated is responsible for Best's macular dystrophy (BMD).
cc Polynucleotides encoding GC1E are useful for diagnosing whether a
cc patient carries a mutation in the GC1E gene. Normal and mutated
cc GC1E proteins are useful for identifying activators and/or inhibitors
cc of these proteins, in order to treat BMD. The GC1E gene offers a
cc simpler and cheaper method of diagnosing BMD without the need for the
cc presence of the patient. The gene may also be useful to discovering
cc the genetic cause of age-related macular dystrophy.
xx

SQ Sequence 16125 BP; 3988 A; 4175 C; 4215 G; 3726 T; 21 other;

| Query Match | Best Local Similarity | Score 4990: | DB 20: | Length 16125: |
|---------------|--|-------------|--------------|---------------------------------|
| Matches 5000: | Conservative | 100.0%; | Pred. No. 0; | Mismatches 0; Indels 0; Gaps 0; |
| OY 1 | CCAAAAAATTTGTCCTTGGGGGTGGGGGCACAAAGCGGAAGAGGGCATTTTGAGCA | 60 | | |
| Db 1 | CCAAAAAATTTGTCCTTGGGGGTGGGGGCACAAAGCGGAAGAGGGCATTTTGAGCA | 60 | | |
| OY 61 | AATTGGCTTAATGCCACCAAGGGCTTTAACACTTAGGTGGTGGTTACAGGTGCA | 120 | | |
| Db 61 | AATTGGCTTAATGCCACCAAGGGCTTTAACACTTAGGTGGTGGTTACAGGTGCA | 120 | | |
| OY 121 | GGAACAACCACATGGCACAGATATCTATGTAAACCACTGCACATCATGTATACCTA | 180 | | |
| Db 121 | GGAACAACCACATGGCACAGATATCTATGTAAACCACTGCACATCATGTATACCTA | 180 | | |
| OY 181 | TGTAAACAACCTGGTACATTTCTGCACACGTATCCAGAGCTTTAGAGTGAATAAAAGT | 240 | | |
| Db 181 | TGTAAACAACCTGGTACATTTCTGCACACGTATCCAGAGCTTTAGAGTGAATAAAAGT | 240 | | |
| OY 241 | GSTGTGTGAATAAATCACCCTGAATCTGACATTAAGTTCAGTTCATTTCCAGAGG | 300 | | |
| Db 241 | GSTGTGTGAATAAATCACCCTGAATCTGACATTAAGTTCAGTTCATTTCCAGAGG | 300 | | |
| OY 301 | AGAGGGTACAGGAAGGAGGATAGAGTGGGTTAAGACACAAGTCATTTATTAATA | 360 | | |
| Db 301 | AGAGGGTACAGGAAGGAGGATAGAGTGGGTTAAGACACAAGTCATTTATTAATA | 360 | | |
| OY 361 | TCAGGGCTTCGGAAAGTTAGTCCCAAAACCAACACATCTCATTAATCCCTGCAGTCTTG | 420 | | |
| Db 361 | TCAGGGCTTCGGAAAGTTAGTCCCAAAACCAACACATCTCATTAATCCCTGCAGTCTTG | 420 | | |
| OY 421 | ATTAAATGAACAATCCCTAAGGCCACAGACTCAACATCTGGAGAAGAAATCCAGAAACT | 480 | | |
| Db 421 | ATTAAATGAACAATCCCTAAGGCCACAGACTCAACATCTGGAGAAGAAATCCAGAAACT | 480 | | |
| OY 481 | GCCCCCTTAATTAACATTTGGGCGATTTTACGGGCTCTTAAGAACCAAGAACCACTGCTG | 540 | | |
| Db 481 | GCCCCCTTAATTAACATTTGGGCGATTTTACGGGCTCTTAAGAACCAAGAACCACTGCTG | 540 | | |
| OY 541 | CCTAGAGCTTGCTCTTTCATTGAACAATACAGAAGAGTGTAGTGAACACCCAC | 600 | | |
| Db 541 | CCTAGAGCTTGCTCTTTCATTGAACAATACAGAAGAGTGTAGTGAACACCCAC | 600 | | |
| OY 601 | ACTTCCAACAGCTTGAGAGAGCCCTTGAGTATGATATGATTAATTAATTTGAATC | 660 | | |
| Db 601 | ACTTCCAACAGCTTGAGAGAGCCCTTGAGTATGATATGATTAATTAATTTGAATC | 660 | | |
| OY 661 | ACATGCTGAGATTTTACACACAGCTGCCCTGGGGAGTCTGGGCAATTAATTCATATGAC | 720 | | |
| Db 661 | ACATGCTGAGATTTTACACACAGCTGCCCTGGGGAGTCTGGGCAATTAATTCATATGAC | 720 | | |
| OY 721 | TGGCTGGGTGAAGGACAGACAGATTAATCCAGGGCTTCTGTCATAACCCACAGACT | 780 | | |
| Db 721 | TGGCTGGGTGAAGGACAGACAGATTAATCCAGGGCTTCTGTCATAACCCACAGACT | 780 | | |
| OY 781 | CACCCCTTCACACGCCCCGAGGCTTCTTCATCTTCATCTCTGAAGCAACTTACGA | 840 | | |
| Db 781 | CACCCCTTCACACGCCCCGAGGCTTCTTCATCTTCATCTCTGAAGCAACTTACGA | 840 | | |

Db 781 CACCCCTCCACACAGCCCGGAGGCTTCCTTCATCTCTGAAACAATTACTGA 840
OY 841 TGGGCTCCGACCAATACACAGCCAGAAATACATATGATGACACAGCAAGCAATACGA 900
Db 841 TGGGCTCCGACCAATACACAGCCAGAAATACATATGATGACACAGCAAGCAATACGA 900
OY 901 GCTCCCTGACATATGCAAAATTCCTCATTTTACTAGGGTGAATGAAATCCCAAGCA 960
Db 901 GCTCCCTGACATATGCAAAATTCCTCATTTTACTAGGGTGAATGAAATCCCAAGCA 960
OY 961 ACACATCTTTTACATTAAGGAGGACCTGAGGCTGAGAGAGAGCTGAATCCCGGGG 1020
Db 961 ACACATCTTTTACATTAAGGAGGACCTGAGGCTGAGAGAGAGCTGAATCCCGGGG 1020
OY 1021 TCACACACACAGGTGGCAAGGCTGGACAGAAACAGAGACTGTGACTGCAGCCGGT 1080
Db 1021 TCACACACACAGGTGGCAAGGCTGGACAGAAACAGAGACTGTGACTGCAGCCGGT 1080
OY 1081 ATTCATCTTTTCCATAGCCACAGGCTGTCAAAAGACCCAGGCGCTAGTCAGAGGCTCC 1140
Db 1081 ATTCATCTTTTCCATAGCCACAGGCTGTCAAAAGACCCAGGCGCTAGTCAGAGGCTCC 1140
OY 1141 TCCCTCTGAGAGATCTGCGACAGAGATTGAAGCTCAGACAGCCCTTAACCCCGCA 1200
Db 1141 TCCCTCTGAGAGATCTGCGACAGAGATTGAAGCTCAGACAGCCCTTAACCCCGCA 1200
OY 1201 CTCTCTGCAAGGCTCAGAGGGGTGAGAACACTGCTGAGCAGATCCTTTAGCCTGGA 1260
Db 1201 CTCTCTGCAAGGCTCAGAGGGGTGAGAACACTGCTGAGCAGATCCTTTAGCCTGGA 1260
OY 1261 TTTTAGGCGCATGCTAGAGGGGGGTGTTCCCTTAATTCAGCCCTGTCTCAGCCCA 1320
Db 1261 TTTTAGGCGCATGCTAGAGGGGGGTGTTCCCTTAATTCAGCCCTGTCTCAGCCCA 1320
OY 1321 CCTTCCAGAGAAATTAAGAGGGGCCAATGGCCAGGCTGTGCTAGCCGCTTCTGAGA 1380
Db 1321 CCTTCCAGAGAAATTAAGAGGGGCCAATGGCCAGGCTGTGCTAGCCGCTTCTGAGA 1380
OY 1381 GATTACAGAGAGGACTTAAGACAGAGACTCCTTTGTGAGAGTCTGTGCTTAAAGAGTCAA 1440
Db 1381 GATTACAGAGAGGACTTAAGACAGAGACTCCTTTGTGAGAGTCTGTGCTTAAAGAGTCAA 1440
OY 1441 GTGAGCGGCTCAGCACTACAGTGGGAGTGGCCAGCTCTTAAGAGTGGGAGGCACT 1500
Db 1441 GTGAGCGGCTCAGCACTACAGTGGGAGTGGCCAGCTCTTAAGAGTGGGAGGCACT 1500
OY 1501 GGGCCAGAGTCCCGAGGAGTCCGACAGCTAGTCCGACAGACTTCTGTGGATCATCG 1560
Db 1501 GGGCCAGAGTCCCGAGGAGTCCGACAGCTAGTCCGACAGACTTCTGTGGATCATCG 1560
OY 1561 GACCCACCTGGAAACCCACCTGTGAGTACAAAGTCCGAGGTGAGTGGGCTT 1620
Db 1561 GACCCACCTGGAAACCCACCTGTGAGTACAAAGTCCGAGGTGAGTGGGCTT 1620
OY 1621 TGAGGCTTCAAGGTTGATGAGGCCATCTTGGCTATTTGTGAGGATATGACACACAGGC 1680
Db 1621 TGAGGCTTCAAGGTTGATGAGGCCATCTTGGCTATTTGTGAGGATATGACACACAGGC 1680
OY 1681 AGCAGATGGCAGGTTGTGGGACCTGTGTCTGTGCAAAATGCGCTGAGGTGGAAATG 1740
Db 1681 AGCAGATGGCAGGTTGTGGGACCTGTGTCTGTGCAAAATGCGCTGAGGTGGAAATG 1740
OY 1741 AGCTTGTGTGATCAGAGGACAGACAGGCTGAGGCTGAGCAAAACACAGGGA 1800
Db 1741 AGCTTGTGTGATCAGAGGACAGACAGGCTGAGGCTGAGCAAAACACAGGGA 1800
OY 1801 AAGATGAGAGGGGCTCATCATCTACATCAAAATTAATTAATAGAGTCCCTTAATAAATA 1860
Db 1801 AAGATGAGAGGGGCTCATCATCTACATCAAAATTAATTAATAGAGTCCCTTAATAAATA 1860
OY 1861 GAAGGTCTCTCTTGTGATAGAAAGAGGAGAGGGGTTGTCTTAATAATATAGG 1920
Db 1861 GAAGGTCTCTCTTGTGATAGAAAGAGGAGAGGGGTTGTCTTAATAATATAGG 1920

OY 1921 GAGAGCCGCCCTCTCAAAAATTAAGGAGGAGGAGACCCAAAGCCCGGAGTGTGTGTT 1980
Db 1921 GAGAGCCGCCCTCTCAAAAATTAAGGAGGAGGAGACCCAAAGCCCGGAGTGTGTGTT 1980
OY 1981 TTCCAGGGGGAGCTCGAACCTTTAGAGGAGCGGTGGAGAACCGCTGTATTCAGGCTC 2040
Db 1981 TTCCAGGGGGAGCTCGAACCTTTAGAGGAGCGGTGGAGAACCGCTGTATTCAGGCTC 2040
OY 2041 TCGAGAGAAAAGAGAGCGGCCCGCAAAATATCCCTCCCGGCGATTAAGAAATGTGTGC 2100
Db 2041 TCGAGAGAAAAGAGAGCGGCCCGCAAAATATCCCTCCCGGCGATTAAGAAATGTGTGC 2100
OY 2101 CTCTCTCAAAAAGATGAAGAGGAGACCGAGTGTATGTGTATTTTAAACTCCA 2160
Db 2101 CTCTCTCAAAAAGATGAAGAGGAGACCGAGTGTATGTGTATTTTAAACTCCA 2160
OY 2161 GGTAGNNNNNNNNNTGCTTCAATTAATTTTATTTAGGCGCTTCTAGAGAACAMGA 2220
Db 2161 GGTAGNNNNNNNNNTGCTTCAATTAATTTTATTTAGGCGCTTCTAGAGAACAMGA 2220
OY 2221 GGAGCTTCCATCTGAGAGAGAAACAGGAGAAACAGGAGCAATATCTGTATTAATTCA 2280
Db 2221 GGAGCTTCCATCTGAGAGAGAAACAGGAGAAACAGGAGCAATATCTGTATTAATTCA 2280
OY 2281 AGTAGATTAAGTCTCTAGAAAATATCAAGCAAGGTAGAGAGACACAGACCCGGTG 2340
Db 2281 AGTAGATTAAGTCTCTAGAAAATATCAAGCAAGGTAGAGAGACACAGACCCGGTG 2340
OY 2341 GCAGTGGGCTATTTCCAGGTTGGATGTGGAAACATCTTTTAAAGGAACTG 2400
Db 2341 GCAGTGGGCTATTTCCAGGTTGGATGTGGAAACATCTTTTAAAGGAACTG 2400
OY 2401 AGTGGAGAAAGAACCAAGTATCTCAGAGAGCTTCTCAGAGAGAGATCAGCA 2460
Db 2401 AGTGGAGAAAGAACCAAGTATCTCAGAGAGCTTCTCAGAGAGAGATCAGCA 2460
OY 2461 GGTGAGAAAGCCCTGAGACCACTTCAATTAATTTAGCATCTTACCAGCTAG 2520
Db 2461 GGTGAGAAAGCCCTGAGACCACTTCAATTAATTTAGCATCTTACCAGCTAG 2520
OY 2521 GTTCCATTAATGGAATATGAGTATGAGTGGAGAGGCTGCTGCTTCCATCTT 2580
Db 2521 GTTCCATTAATGGAATATGAGTATGAGTGGAGAGGCTGCTGCTTCCATCTT 2580
OY 2581 CTGACACTAGGAGTGTGAGAGAGCTTGGAGGCTTAACGAACAGATGGCTGAGAACACT 2640
Db 2581 CTGACACTAGGAGTGTGAGAGAGCTTGGAGGCTTAACGAACAGATGGCTGAGAACACT 2640
OY 2641 GCTAGGCCAGAGAGACTGAGCTTATGAGTATGACATTTGCTGCTTACTGCTTGTGCG 2700
Db 2641 GCTAGGCCAGAGAGACTGAGCTTATGAGTATGACATTTGCTGCTTACTGCTTGTGCG 2700
OY 2701 TTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2760
Db 2701 TTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2760
OY 2761 GCTTGAAGTCAATGAGGATGATGAGTCACTGCACTGCACTGCACTGCTGCTGCTGCA 2820
Db 2761 GCTTGAAGTCAATGAGGATGATGAGTCACTGCACTGCACTGCACTGCTGCTGCTGCA 2820
OY 2821 TTCTCTGCTCAGGCTCTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 2880
Db 2821 TTCTCTGCTCAGGCTCTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 2880
OY 2881 ATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2940
Db 2881 ATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2940
OY 2941 TCCCTGACCTTATGAGTATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000
Db 2941 TCCCTGACCTTATGAGTATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000

3001 GCCACTGCCCCAGTATTATAGAAAGTTAAAGCAGCATGCGAATGACACAGCCCTATCTA 3060
3001 GCCACTGCCCCAGTATTATAGAAAGTTAAAGCAGCATGCGAATGACACAGCCCTATCTA 3060
3061 GCTTTCCCTGCGCAAGCAAGGAGGAGCGCTGGGCTCAGTTTCTGCGTTTCTACTTCC 3120
3061 GCTTTCCCTGCGCAAGCAAGGAGGAGCGCTGGGCTCAGTTTCTGCGTTTCTACTTCC 3120
3121 AAAAGGACATCAGAACTGGCAGGCGCTTGGAGACACTTCATCCACTCTAGGTCCTCT 3180
3121 AAAAGGACATCAGAACTGGCAGGCGCTTGGAGACACTTCATCCACTCTAGGTCCTCT 3180
3121 AAAAGGACATCAGAACTGGCAGGCGCTTGGAGACACTTCATCCACTCTAGGTCCTCT 3180
3181 ATGGAGAGTTGAGTCCAGAGAGGAGGCTCAGACAGGCTTGACAGGCGCTCTG 3240
3181 ATGGAGAGTTGAGTCCAGAGAGGAGGCTCAGACAGGCTTGACAGGCGCTCTG 3240
3241 ATCCCTACAAACCCCAATCGGTGTCCTCTACAGAGACCCCAAGCCCACTGCTGAG 3300
3241 ATCCCTACAAACCCCAATCGGTGTCCTCTACAGAGACCCCAAGCCCACTGCTGAG 3300
3301 CCCACTGCTGGCATGACCATCTTACACAAAGCCAGTGGCTTAATGCCCTTAAAGCT 3360
3301 CCCACTGCTGGCATGACCATCTTACACAAAGCCAGTGGCTTAATGCCCTTAAAGCT 3360
3301 CCCACTGCTGGCATGACCATCTTACACAAAGCCAGTGGCTTAATGCCCTTAAAGCT 3360
3361 CCTTCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
3361 CCTTCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
3361 CCTTCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
3421 TCCATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
3421 TCCATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
3481 GGTGGGCGCGGGGCGCTGGGAGAGTGTGGCTGGGCTGGGAGCTGGGAGCTCTGGGG 3540
3481 GGTGGGCGCGGGGCGCTGGGAGAGTGTGGCTGGGCTGGGAGCTGGGAGCTCTGGGG 3540
3481 GGTGGGCGCGGGGCGCTGGGAGAGTGTGGCTGGGCTGGGAGCTGGGAGCTCTGGGG 3540
3541 CCTCCAGCAGCTCAGGCGCCAGTGCACAGTGCACAGTGCACAGTGCACAGTGCACAG 3600
3541 CCTCCAGCAGCTCAGGCGCCAGTGCACAGTGCACAGTGCACAGTGCACAGTGCACAG 3600
3541 CCTCCAGCAGCTCAGGCGCCAGTGCACAGTGCACAGTGCACAGTGCACAGTGCACAG 3600
3541 CCTCCAGCAGCTCAGGCGCCAGTGCACAGTGCACAGTGCACAGTGCACAGTGCACAG 3600
3601 ACCAGCTCTGGGCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 3660
3601 ACCAGCTCTGGGCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 3660
3661 CACACTGAGAGGCTGCTCAAGCCAGGCGGAGGCTTAAAGCAAGGAGGAGAGAGAG 3720
3661 CACACTGAGAGGCTGCTCAAGCCAGGCGGAGGCTTAAAGCAAGGAGGAGAGAGAG 3720
3661 CACACTGAGAGGCTGCTCAAGCCAGGCGGAGGCTTAAAGCAAGGAGGAGAGAGAG 3720
3721 AGGAGAGCCCTGAGGCGGAGGCTTGTGCGGCCCACTTACTGCGCTGTTTACTGAA 3780
3721 AGGAGAGCCCTGAGGCGGAGGCTTGTGCGGCCCACTTACTGCGCTGTTTACTGAA 3780
3721 AGGAGAGCCCTGAGGCGGAGGCTTGTGCGGCCCACTTACTGCGCTGTTTACTGAA 3780
3781 TCCACACAGACTCATAGGCGCCAGTATTAAGGAGAGAGAGAGAGAGAGAGAGAG 3840
3781 TCCACACAGACTCATAGGCGCCAGTATTAAGGAGAGAGAGAGAGAGAGAGAGAGAG 3840
3781 TCCACACAGACTCATAGGCGCCAGTATTAAGGAGAGAGAGAGAGAGAGAGAGAGAG 3840
3841 GAGAGAGATGAGAGTCTCACTGTGTCCAGGCTGCTGCAACTCTTACAGGCTCAAGCA 3900
3841 GAGAGAGATGAGAGTCTCACTGTGTCCAGGCTGCTGCAACTCTTACAGGCTCAAGCA 3900
3841 GAGAGAGATGAGAGTCTCACTGTGTCCAGGCTGCTGCAACTCTTACAGGCTCAAGCA 3900
3901 TCCCGCTGCTTACCTTCCCAAGGAGGCTGAGTATTAAGGAGTATTAAGGAGTATTAAG 3960
3901 TCCCGCTGCTTACCTTCCCAAGGAGGCTGAGTATTAAGGAGTATTAAGGAGTATTAAG 3960
3901 TCCCGCTGCTTACCTTCCCAAGGAGGCTGAGTATTAAGGAGTATTAAGGAGTATTAAG 3960
3961 AACACATGTAATTTTTTTTTTTTTTTTTTTTTTTTTTGAAGAGGTTTCACTCATCAC 4020
3961 AACACATGTAATTTTTTTTTTTTTTTTTTTTTTTTTTGAAGAGGTTTCACTCATCAC 4020
3961 AACACATGTAATTTTTTTTTTTTTTTTTTTTTTTTTTGAAGAGGTTTCACTCATCAC 4020
4021 CAGGCTGAGTGTGAGGAGGAGATTTGGCTCAGTGAACCTCTGCTCCAGGTCAG 4080
4021 CAGGCTGAGTGTGAGGAGGAGATTTGGCTCAGTGAACCTCTGCTCCAGGTCAG 4080
4021 CAGGCTGAGTGTGAGGAGGAGATTTGGCTCAGTGAACCTCTGCTCCAGGTCAG 4080
4081 GCGATTCTCTGCTTACGCTCTGAGTGTGAGTATTAAGGAGAGAGAGAGAGAGAGAG 4140

4081 GCGATTCTCTGCTTACGCTCTGAGTACTGGAATTAAGGAGAGAGAGAGAGAGAGAG 4140
4141 GCGTAATTTTTTTTTTTTTTTTTTTTTTTTTTGAAGAGAGAGGTTTCACTGTTGGCAG 4200
4141 GCGTAATTTTTTTTTTTTTTTTTTTTTTTTTTGAAGAGAGAGGTTTCACTGTTGGCAG 4200
4141 GCGTAATTTTTTTTTTTTTTTTTTTTTTTTTTGAAGAGAGAGGTTTCACTGTTGGCAG 4200
4201 GCTGCTTGAACCCCTGAGGAGTCAAGGAGTCAAGGAGTCAAGGAGTCAAGGAGTCAAG 4260
4201 GCTGCTTGAACCCCTGAGGAGTCAAGGAGTCAAGGAGTCAAGGAGTCAAGGAGTCAAG 4260
4201 GCTGCTTGAACCCCTGAGGAGTCAAGGAGTCAAGGAGTCAAGGAGTCAAGGAGTCAAG 4260
4261 ATTAAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4320
4261 ATTAAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4320
4261 ATTAAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4320
4321 ATTAAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4380
4321 ATTAAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4380
4321 ATTAAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4380
4381 GGGGCGAAGGTGGGAGTCACTTGAAGGAGTCACTGAGGAGTCACTGAGGAGTCACTGAG 4440
4381 GGGGCGAAGGTGGGAGTCACTTGAAGGAGTCACTGAGGAGTCACTGAGGAGTCACTGAG 4440
4381 GGGGCGAAGGTGGGAGTCACTTGAAGGAGTCACTGAGGAGTCACTGAGGAGTCACTGAG 4440
4441 CCCGCTCTACCAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 4500
4441 CCCGCTCTACCAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 4500
4441 CCCGCTCTACCAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 4500
4501 AGCTACTTGGAGAGTGGAGGAGTGGAGGAGTGGAGGAGTGGAGGAGTGGAGGAGTGGAG 4560
4501 AGCTACTTGGAGAGTGGAGGAGTGGAGGAGTGGAGGAGTGGAGGAGTGGAGGAGTGGAG 4560
4501 AGCTACTTGGAGAGTGGAGGAGTGGAGGAGTGGAGGAGTGGAGGAGTGGAGGAGTGGAG 4560
4561 AGCTATGATACACACCTGAGTCACTGAGGAGTCACTGAGGAGTCACTGAGGAGTCACTGAG 4620
4561 AGCTATGATACACACCTGAGTCACTGAGGAGTCACTGAGGAGTCACTGAGGAGTCACTGAG 4620
4561 AGCTATGATACACACCTGAGTCACTGAGGAGTCACTGAGGAGTCACTGAGGAGTCACTGAG 4620
4621 ATATGTTTATCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4680
4621 ATATGTTTATCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4680
4621 ATATGTTTATCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4680
4681 AAAAGCATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 4740
4681 AAAAGCATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 4740
4681 AAAAGCATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 4740
4741 AATAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4800
4741 AATAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4800
4741 AATAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4800
4801 AGTTAGGCGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4860
4801 AGTTAGGCGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4860
4801 AGTTAGGCGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4860
4861 GAGAGATCACTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4920
4861 GAGAGATCACTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4920
4861 GAGAGATCACTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4920
4921 TCTCTATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAAT 4980
4921 TCTCTATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAAT 4980
4921 TCTCTATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAAT 4980
4981 TTTGGAGGCGGAGGCGGT 5000
4981 TTTGGAGGCGGAGGCGGT 5000
4981 TTTGGAGGCGGAGGCGGT 5000

RESULT 2
ABAI4559
ID ABAI4559 standard; DNA: 16650 BP.
XX ABAI4559;
AC 23-JAN-2002 (first entry)
DT 23-JAN-2002 (first entry)
XX Human nervous system related polynucleotide seq ID NO 6890.

XX Human; nootropic; neuroprotective; cytosstatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerrary;
KW antiparkinsonian; antischizling; antianaemic; antiarthritic; cancer;
KW antileumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antidiabetic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
OS Homo sapiens.
XX MO200159063-A2.
PN 16-AUG-2001.
XX 17-JAN-2001; 2001WO-US01334.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0230963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225275.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249247.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.

| | | | | |
|----|------|--|---|--|
| | XX | PA | (HUMA-) HUMAN GENOME SCI INC. | |
| | XX | PI | Rosen CA, Barash SC, Ruben SM; | |
| | DR | WIPI: | 2001-541565/60. | |
| | XX | PT | Nucleic acids encoding 3224 human nervous system antigen polypeptides, | |
| | XX | PT | useful for preventing, diagnosing and/or treating nervous system | |
| | XX | PT | cancers and metastases - | |
| | PS | Disclosure: | SEQ ID NO 6890: 1701bp + Sequence Listing; English. | |
| | CC | XX | The invention relates to novel genes (ABAI1004-ABA23534) and proteins | |
| | CC | XX | (ABAI14678-ABBI8001) useful for preventing, treating or ameliorating | |
| | CC | XX | medical conditions e.g. by protein or gene therapy. The genes are | |
| | CC | XX | isolated from a range of human tissues disclosed in the specification. | |
| | CC | XX | The nucleic acids, proteins, antibodies and (ant)agonists are useful | |
| | CC | XX | in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast | |
| | CC | XX | and ovarian cancer and other cancers of the adrenal gland, bone, bone | |
| | CC | XX | marrow, breast, gastrointestinal tract, liver, lung, or urogenital; | |
| | CC | XX | (b) immune disorders e.g. Addison's disease, diabetes mellitus, Crohn's | |
| | CC | XX | disease, hemolytic anaemia, autoimmune thyroiditis, allergies, autoimmune | |
| | CC | XX | diseases, multiple sclerosis, rheumatoid arthritis and ulcerative | |
| | CC | XX | colitis; (c) cardiovascular disorders such as myocardial ischaemias; | |
| | CC | XX | (d) wound healing; (e) neurological diseases such as viral, bacterial, fungal | |
| | CC | XX | epilepsy; and (f) infectious diseases such as viral, bacterial, fungal | |
| | CC | XX | and parasitic infections. | |
| | CC | XX | Note: The sequence data for this patent did not form part of the | |
| | CC | XX | printed specification, but was obtained in electronic format directly | |
| | CC | XX | from WIPO at ftp.wipo.int/pub/published_pcr_sequences. | |
| | SO | Sequence | 16650 BP; 4168 A; 4271 C; 4308 G; 3903 T; 0 other; | |
| | | Query Match | 55.2%; Score 2758.2; DB 22; Length 16650; | |
| | | Best Local Similarity | 98.4%; Pred. No. 0; Mismatches 38; Indels 8; Gaps 2; | |
| | | Matches 2803; Conservative | 0; Mismatches 38; Indels 8; Gaps 2; | |
| OY | 2152 | AAAACTCCAGTAGNNNNNNNNNTGCTTCACGTAAATTTTATTGAGCGCCTTACGAG | 2211 | |
| DB | 285 | AAGANTGGAGGGGCATCAATCACGCTTCAGTAATTTTATTGAGCGCCTTACGAG | 344 | |
| OY | 2212 | AACACANAGAGACTTCATTTTGAGAGGAACAACGAGAACAGGCACATNATCCGT | 2271 | |
| DB | 345 | AACACAMAGGAGGCTTCATTTTGAGAGGAACAACGAGAACAGGCACATNATCCGT | 404 | |
| OY | 2272 | ATAAATTCAGATGATAGTATGCTCTCAGAAATATCAAGCAAGTATGAGACACAGA | 2331 | |
| DB | 405 | ATAAATTCAGATGATAGTATGCTCTCAGAAATATCAAGCAAGTATGAGACACAGA | 464 | |
| OY | 2332 | GCAACGGTGGCAGTGGGGCTCATTTCCAGGTGGATGGTGGAACTCTTTTAAAG | 2391 | |
| DB | 465 | GCAACGGTGGCAGTGGGGCTCATTTCCAGGTGGATGGTGGAACTCTTTTAAAG | 524 | |
| OY | 2392 | GGAACCTGGAGTGGGAAGAACCATGACAGTATTCACAGAAAGAGCTCTCCAGCAGGA | 2451 | |
| DB | 525 | GGAACCTGGAGTGGGAAGAACCATGACAGTATTCACAGAAAGAGCTCTCCAGCAGGA | 584 | |
| OY | 2452 | AGATTCAGACAGGTGGAAAGGCCCTGAGACCACATTCAGTAAACATCATTTGACATCTCT | 2511 | |
| DB | 585 | AGATTCAGACAGGTGGAAAGGCCCTGAGACCACATTCAGTAAACATCATTTGACATCTCT | 644 | |
| OY | 2512 | ACACAGTAGTTTCATTTAGGAATGGGAATATGATGGTGGACAGAGGCGTGGGAGCCT | 2571 | |
| DB | 645 | ACACAGTAGTTTCATTTAGGAATGGGAATATGATGGTGGACAGAGGCGTGGGAGCCT | 704 | |

[illegible]

1777 AAAAAAGTCAGACTAGAGGCTGCTCAAGCCAGAGCCAGAGGCTTTTACCCACCCCTCC 1836
1837 TCCACCCCGAGAGAGACCCCTGAGCCAGCCAGCTTTGCTCCTCCACTCAGGCGCTGT 3771
1897 TCCACCCCGAGAGAGACCCCTGAGCCAGCCAGCTTTGCTCCTCCACTCAGGCGCTGT 1896
3772 TTTACTGAATCCACACAGACTATAGAGCCACATATAGATTTAAAAAGAGAGAGAG 3831
1897 TTTACTGAATCCACACAGACTATAGAGCCACATATAGATTTAAAAAGAGAGAGAG 1956
3832 AGAGAGAGAGAGAGAGATGAGTCTCAGTCTGTTGTCAGAGCTGCTCAGTCTCAG 3891
1957 AGAGAGAGAGAGAGAGATGAGTCTCAGTCTGTTGTCAGAGCTGCTCAGTCTCAG 2016
3892 CTCAGCAATCCCTCCCTTACGCTCCTCAGGAGGCTGAGTATACAGTCTGAGCTCTG 3951
2017 CTCAGCAATCCCTCCCTTACGCTCCTCAGGAGGCTGAGTATACAGTCTGAGCTCTG 2076
3952 CACTTGACCAACACATGATGATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 4011
2077 CACTTGACCAACACATGATGATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2136
4012 TCCATCACCCAGGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 4071
2137 TCCATCACCCAGGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 2196
4072 CAGGTGAGAGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4131
2197 CAGGTGAGAGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2256
4132 ACCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4191
2257 ACCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2316
4192 GTTGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4251
2317 GTTGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2376
4252 AGTCTGGAGTATACAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 4311
2377 AGTCTGGAGTATACAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 2436
4312 TATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4371
2437 TATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2496
4372 GCACCTTGAAGGCGCAGGCTGAGGCGCAGGCTGAGGCGCAGGCTGAGGCGCAGGCTGAG 4431
2497 GCACCTTGAAGGCGCAGGCTGAGGCGCAGGCTGAGGCGCAGGCTGAGGCGCAGGCTGAG 2556
4432 TAGTGAAGACCCGCTCTACCAAAATTTAAATAATTAATTAATTAATTAATTAATTAATTA 4491
2557 TAGTGAAGACCCGCTCTACCAAAATTTAAATAATTAATTAATTAATTAATTAATTAATTA 2616
4492 TGTGTCTCCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 4551
2617 TGTGTCTCCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 2676
4552 GCTGAGAGAGTATAGTATACACACAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 4611
2677 GCTGAGAGTATAGTATACACACAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 2736
4612 GCAAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4671
2737 GCAAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2796
4672 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4731
2797 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2856
4732 TGTGTCTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4791

2857 TGTGTCTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2916
4792 TTTAAAAAAGTTAGGCGCAGCCAGAGGCTCAGACCTGTATATCCAGACATTTGGAG 4851
2917 TTTAAAAAAGTTAGGCGCAGCCAGAGGCTCAGACCTGTATATCCAGACATTTGGAG 2976
4852 GCCAAGCAGAGAGATCAGTCTGAGCCAGAGAGTTAGGAGCTTGGAGCAACATAGGAGATC 4911
2977 GCCAAGCAGAGAGATCAGTCTGAGCCAGAGAGTTAGGAGCTTGGAGCAACATAGGAGATC 3036
4912 CTGATCTGCTCTATATAAATAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 4971
3037 CTGATCTGCTCTATATAAATAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 3096
4972 TCCAGACATTTGGAGCGCGAGCGCGGT 5000
3097 TCCAGACATTTGGAGCGCGAGCGCGGT 3125

RESULT 3
ABAI4557
ID ABAI4557 standard; DNA; 18530 BP.
XX
AC ABAI4557;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6888.
XX
KW Human; neotropic; neuroprotective; cytosolic; dermatological; vitruide;
KW immunosuppressive; anti-inflamatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antischilling; antianemic; antiallergic; cancer;
KW antileukemic; hepatocytic; cerebroprotective; antineoplastic;
KW antiallergic; antidiabetic; antileuc; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN W0200159063-A2.
XX
PD 16-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0198076.
PR 17-MAR-2000; 2000US-0198076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0218880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225457.

2152 AAAACCTCAGTAGNNNNNNNTGCTCAGTAATTTTATTTAGAGGCTTCTAGAG 2211
285 AAGATGAGAGGGGATCATCTAGCTGCTCAATTAATTTTATTTAGAGGCTTCTAGAG 344
2212 AACACAGAGAGGCTTCATCTGAGAGGAAACAGGAGGAAACAGGAGATATCTGT 2271
345 AACACAGAGAGGCTTCATCTGAGAGGAAACAGGAGGAAACAGGAGATATCTGT 404
2272 ATAAATTCAGTAGTAGTAGTCTCTCTAGAAATATCAACAGGATAGAGACACAGA 2331
405 ATAAATTCAGTAGTAGTAGTCTCTCTAGAAATATCAACAGGATAGAGACACAGA 464
2332 GCACCGGTGAGAGGGGCTCATTTTCCAGGTGATGTTGGGAAATCCTTTCTAAAG 2391
465 GCACCGGTGAGAGGGGCTCATTTTCCAGGTGATGTTGGGAAATCCTTTCTAAAG 524
2392 GGAACCTGAGATGGGAAAGAAACATGAGATATCTCAGAGAAAGCTTCTCCAGGACGA 2451
525 GGAACCTGAGATGGGAAAGAAACATGAGATATCTCAGAGAAAGCTTCTCCAGGACGA 584
2452 AGATCAGAGGTGGAAGGAGGCTGAGGACCATTCAGTAAACATTTAGATCTCT 2511
585 AGATCAGAGGTGGAAGGAGGCTGAGGACCATTCAGTAAACATTTAGATCTCT 644
2512 ACCAGCTAGGTTCCATTTATGGAAATGGAAATGGTGGTGGACAGGCTGCTGCTCT 2571
645 ACCAGCTAGGTTCCATTTATGGAAATGGAAATGGTGGTGGACAGGCTGCTGCTCT 704
2572 TCCATACCTTCACACTAGAGGTGTTGAGAGCTTTGGAGCTTAACGAACAGATGGGT 2631
705 TCCATACCTTCACACTAGAGGTGTTGAGAGCTTTGGAGCTTAACGAACAGATGGGT 764
2632 GGAACACAGGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2691
765 GGAACACAGGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 824
2692 CCTTGTGCTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2751
825 CCTTGTGCTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 884
2752 CTTACCCAGGCTGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2811
885 CTTACCCAGGCTGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 944
2812 ATCAACGATTCCTGCTCAGCTCAGCTCAGTGAAGTGGATTAACAGGACCCGACAC 2871
945 TTCAACGATTCCTGCTCAGCTCAGCTCAGTGAAGTGGATTAACAGGACCCGACAC 1004
2872 GCCGTGATTAATTTTGTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2931
1005 GCCGTGATTAATTTTGTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1064
2932 GTTCTGAACTCCGAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2991
1065 GTTCTGAACTCCGAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1124
2992 TAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3051
1125 TAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1184
3052 GCCGTATGATCTTCCCTGCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3111
1185 GCCGTATGATCTTCCCTGCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1244
3112 TCTACTTCCAAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3171
1245 TCTACTTCCAAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1304
3172 AGGATCCCTATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3231
1305 AGGATCCCTATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1364
3232 GGGGCTGTGATCCCTACAAACCCCAATCGGTGTCCTCTACAGAGACCCAAAGCCAC 3291

1365 GGGGCTGTGATCCCTACAAACCCCAATCGGTGTCCTCTACAGAGACCCAAAGCCAC 1424
3292 CTGCTGAGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3351
1425 CTGCTGAGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1484
3352 GCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3411
1485 GCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1544
3412 ATGCGAGTCCATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3471
1545 ATGCGAGTCCATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1604
3472 GCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3531
1605 GCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1657
3532 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3591
1658 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1717
3592 GGGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3651
1718 GAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1776
3652 TAAAGAGTCACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3711
1777 AAAAGAGTCACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1836
3712 TCCAAACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3771
1837 TCCAAACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1896
3772 TTTAGTGAATCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3831
1897 TTTAGTGAATCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1956
3832 AG 3891
1957 AG 2016
3892 CTCAGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3951
2017 CTCAGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2076
3952 CACTGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4011
2077 CACTGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2136
4012 TCCATACCCAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4071
2137 TCCATACCCAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2196
4072 CAGGTCAAGAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4131
2197 CAGGTCAAGAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2256
4132 ACCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4191
2257 ACCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2316
4192 GTTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4251
2317 GTTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2376
4252 AGTCTGGAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4311
2377 AGTCTGGAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2436
4312 TATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4371

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM-nucleic - nucleic search, using sw model

Run on: June 8, 2003, 18:10:38 ; Search time 317.898 Seconds
(without alignments)
10693.617 Million cell updates/sec

Title: US-09-622-964-4

Perfect score: 2429
Sequence: 1 caggagatccaccagccta.....aaaaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 1244.2 | 51.2 | 1263 | 9 | US-09-746-783-3 |
| 2 | 350.4 | 14.4 | 1326 | 9 | US-10-071-766-10 |
| 3 | 189.2 | 7.8 | 1292 | 9 | US-10-198-846-11070 |
| 4 | 175 | 7.2 | 1198 | 10 | US-09-880-107-2174 |
| 5 | 126.8 | 5.2 | 1350 | 10 | US-09-768-826-16 |
| 6 | 116 | 4.8 | 853 | 9 | US-10-198-846-11246 |
| 7 | 104 | 4.3 | 462 | 9 | US-10-198-846-2561 |
| 8 | 102 | 4.2 | 615 | 9 | US-10-198-846-7843 |
| 9 | 98 | 4.0 | 507 | 9 | US-09-764-872-195 |
| 10 | 90 | 3.7 | 684973 | 10 | US-09-263-959-1 |
| 11 | 89 | 3.7 | 31994 | 9 | US-09-764-904-71 |
| 12 | 89 | 3.7 | 31994 | 9 | US-10-091-548-71 |
| 13 | 89 | 3.7 | 31994 | 9 | US-10-074-095-599 |
| 14 | 89 | 3.7 | 31994 | 10 | US-09-764-860-599 |
| 15 | 88 | 3.6 | 30568 | 9 | US-09-764-877-3851 |
| 16 | 87.8 | 3.6 | 133893 | 9 | US-10-161-510-1 |
| 17 | 87.4 | 3.6 | 4963 | 10 | US-09-764-877-2903 |
| 18 | 87.4 | 3.6 | 23509 | 9 | US-10-291-737-3 |
| 19 | 86.4 | 3.6 | 2625 | 9 | US-09-764-891-8802 |

| | | | | | | |
|----|------|-----|--------|----|---------------------|--------------------|
| 20 | 86.2 | 3.5 | 289 | 9 | US-10-115-278-4 | Sequence 4, Appl1 |
| 21 | 86.2 | 3.5 | 8663 | 9 | US-10-079-854-397 | Sequence 397, App |
| 22 | 86.2 | 3.5 | 8663 | 10 | US-09-764-878-397 | Sequence 397, App |
| 23 | 86.2 | 3.5 | 75899 | 10 | US-09-854-883-243 | Sequence 211, App |
| 24 | 86 | 3.5 | 10894 | 9 | US-10-091-504-2111 | Sequence 2111, App |
| 25 | 86 | 3.5 | 10894 | 10 | US-09-764-869-2111 | Sequence 5805, App |
| 26 | 85.8 | 3.5 | 219 | 9 | US-09-764-891-5805 | Sequence 5805, App |
| 27 | 85.8 | 3.5 | 219 | 9 | US-09-764-891-5806 | Sequence 32677, A |
| 28 | 85.8 | 3.5 | 484 | 9 | US-09-918-995-32677 | Sequence 8703, App |
| 29 | 85.8 | 3.5 | 8018 | 9 | US-09-764-891-8703 | Sequence 3468, App |
| 30 | 85.8 | 3.5 | 16337 | 10 | US-09-764-877-3468 | Sequence 1605, App |
| 31 | 85.8 | 3.5 | 32195 | 9 | US-10-091-504-1605 | Sequence 611, App |
| 32 | 85.8 | 3.5 | 32195 | 9 | US-10-125-540-611 | Sequence 611, App |
| 33 | 85.8 | 3.5 | 32195 | 10 | US-09-764-870-611 | Sequence 617, App |
| 34 | 85.8 | 3.5 | 32195 | 10 | US-09-764-870-617 | Sequence 1605, App |
| 35 | 85.8 | 3.5 | 32195 | 10 | US-09-764-869-1605 | Sequence 6186, App |
| 36 | 85.8 | 3.5 | 32195 | 10 | US-09-764-891-6186 | Sequence 6187, App |
| 37 | 85.6 | 3.5 | 261 | 9 | US-09-764-891-6187 | Sequence 6188, App |
| 38 | 85.6 | 3.5 | 261 | 9 | US-09-764-891-6188 | Sequence 2901, App |
| 39 | 85.6 | 3.5 | 261 | 9 | US-10-198-846-2901 | Sequence 260, App |
| 40 | 85 | 3.5 | 402850 | 9 | US-09-844-653-5 | Sequence 6304, App |
| 41 | 85 | 3.5 | 2581 | 9 | US-09-822-846-260 | Sequence 1373, App |
| 42 | 84.8 | 3.5 | 32191 | 9 | US-09-764-891-6304 | Sequence 1373, App |
| 43 | 84.8 | 3.5 | 184 | 9 | US-09-764-868-1373 | Sequence 2, Appl1 |
| 44 | 84.6 | 3.5 | 288 | 9 | US-10-115-278-2 | |
| 45 | 84.6 | 3.5 | 288 | 9 | US-10-115-278-2 | |

ALIGNMENTS

RESULT 1
US-09-746-783-3
Sequence 3, Appl1
Publication No. US2003044935A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavalie, Edward R.
Racle, Lisa A.
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.
Howes, Steven H.
Fechtner, Kim
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 231
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/746,783
FILING DATE: 21-Dec-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Milasincic, Debra J.
REGISTRATION NUMBER: 46,931
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 742-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1263 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-746-783-3

Query Match 51.2%; Score 1244.2; DB 9; Length 1263;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1246; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
OY 1181 GACAGCTCATAACCCCTTTGGAGAGATGATGATTTTGAAGCAACTGGATTGTC 1240
DB 1 GACAGCTCATAACCCCTTTGGAGAGATGATGATTTTGAAGCAACTGGATTGTC 60
OY 1241 GACAGCAATTTGCAAGTGTCTCCCTGTTGGATGATGATGACACGACCTGCTCGG 1300
DB 61 GACAGCAATTTGCAAGTGTCTCCCTGTTGGATGATGATGACACGACCTGCTCGG 120
OY 1301 ATGAGAGCCGACATGATGTAATTAAGCCGAGCCACAGCCCTTACACAGCTGCTCC 1360
DB 121 ATGAGAGCCGACATGATGTAATTAAGCCGAGCCACAGCCCTTACACAGCTGCTCC 180
OY 1361 GCCAGTTCCTGAGCCTCTTTATGAGCTCCACTTCAACATCAGCCTGGAACAAAGAG 1420
DB 181 GCCAGTTCCTGAGCCTCTTTATGAGCTCCACTTCAACATCAGCCTGGAACAAAGAG 240
OY 1421 GAGATGAGTCCAGCCCAATCAGAGAGAGAGAGATGCTACAGCTGAGCATTTGGC 1480
DB 241 GAGATGAGTCCAGCCCAATCAGAGAGAGAGAGATGCTACAGCTGAGCATTTGGC 300
OY 1481 CGCTTCTTGAAGCTGCAATCCATGATCACCATCTCCCAAGGCAAACTCAAGAGCCAAA 1540
DB 301 CGCTTCTTGAAGCTGCAATCCATGATCACCATCTCCCAAGGCAAACTCAAGAGCCAAA 360
OY 1541 CTACTGTGCCCCAAGAGGAAATCCCTTCTCCAGAGGAGCTGCCCCAAAACACAGAGCA 1600
DB 361 CTACTGTGCCCCAAGAGGAAATCCCTTCTCCAGAGGAGCTGCCCCAAAACACAGAGCA 420
OY 1601 GCCAAACAGAACGTTAGGGCCAGAGAACAAAGAGGCTGGAAGCTTAAAGCTGTGAGC 1660
DB 421 GCCAAACAGAACGTTAGGGCCAGAGAACAAAGAGGCTGGAAGCTTAAAGCTGTGAGC 480
OY 1661 GCCCTCAAGTGTGCCCCACTGTATCAGAGGAGCTTACTACAGTGGCCCCACAGAGGCC 1720
DB 481 GCCCTCAAGTGTGCCCCACTGTATCAGAGGAGCTTACTACAGTGGCCCCACAGAGGCC 540
OY 1721 CTCAGCCCACTCCCATGTTCTTCCCTCCTAGAACATCAGCCTGTCAAAGCTTCAAGT 1780
DB 541 CTCAGCCCACTCCCATGTTCTTCCCTCCTAGAACATCAGCCTGTCAAAGCTTCAAGT 600
OY 1781 GTTCACAGGCAATAGACACCAAGAACAAAGCTTAAAGACTGTGAGTTCTGGGCCAAAGAAA 1840
DB 601 GTTCACAGGCAATAGACACCAAGAACAAAGCTTAAAGACTGTGAGTTCTGGGCCAAAGAAA 660
OY 1841 AGTTTGAATGCTCTCAGAGAGCGATGAGGAGCTTGTGAGAGACCCAGAAATATCTCAA 1900
DB 661 AGTTTGAATGCTCTCAGAGAGCGATGAGGAGCTTGTGAGAGACCCAGAAATATCTCAA 720
OY 1901 GTGAGAGAGAAAACCTGTGAGTTTAACCTGACGAGATATGACAGATATCCCGAAATATCAG 1960
DB 721 GTGAGAGAGAAAACCTGTGAGTTTAACCTGACGAGATATGACAGATATCCCGAAATATCAG 780
OY 1961 CTCAGAAACCTTTGGAACATATCAGCAACCAATCAGACTCACTCAAAAGTTCACAG 2020
DB 781 CTCAGAAACCTTTGGAACATATCAGCAACCAATCAGACTCACTCAAAAGTTCACAG 840
OY 2021 GATCCTTATGAGGCTTGAAGAACAGAGGATGAAGCAATCTTACCTGCTCTCTAATG 2080
DB 841 GATCCTTATGAGGCTTGAAGAACAGAGGATGAAGCAATCTTACCTGCTCTCTAATG 900
OY 2081 GGATGCTTGGCAGCAGGTCTCAGCTGTGTGTACACGACAGAGACATGATCCAGTCA 2140
DB 2141 GGATGCTTGGCAGCAGGTCTCAGCTGTGTGTGTACACGACAGAGACATGATCCAGTCA 2200
```

```
DB 901 GGATGCTTGGCAGCAGGTCTCAGCTGTGTGTGTACACGACAGAGACATGATCCAGTCA 960
OY 2141 CAGCCATACAGCTGTCCACACTGGAAGAACGTTCTTCAACAGCCTGATCAATGTTA 2200
DB 961 CAGCCATACAGCTGTCCACACTGGAAGAACGTTCTTCAACAGCCTGATCAATGTTA 1020
OY 2201 GCTTAATGATTAATAATCCAGACTACTTCAAGCTTTAATGCTTTATTCATAAAACT 2260
DB 1021 GCTTAATGATTAATAATCCAGACTACTTCAAGCTTTAATGCTTTATTCATAAAACT 1080
OY 2261 GTGAAAGCTAGACTGAAACATTTGAAACATTTACTGAGCTCTGAGATTCAGAGTGGGA 2320
DB 1081 GTGAAAGCTAGACTGAAACATTTGAAACATTTACTGAGCTCTGAGATTCAGAGTGGGA 1140
OY 2321 ACCCTTAGTCTTATGATGATCAAGACAGCCACACTTAGTATGATGATGATGATGATG 2380
DB 1141 ACCCTTAGTCTTATGATGATCAAGACAGCCACACTTAGTATGATGATGATGATGATG 1200
OY 2381 GTTAAATATTAATAATACTCGTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 2429
DB 1201 GTTAAATATTAATAATACTCGTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1249
```

RESULT 2
US-10-071-766-10/c
Sequence 10, Application US/10071766
Publication No. US20020192678A1
GENERAL INFORMATION:
APPLICANT: Huel-Mei Chen
TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
FILE REFERENCE: PA-0043 US
CURRENT APPLICATION NUMBER: US/10/071,766
CURRENT FILING DATE: 2002-02-07
NUMBER OF SEQ. ID NOS: 144
SOFTWARE: PERL Program
SEQ ID NO 10
LENGTH: 1326
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
OTHER INFORMATION: Incycle ID No. US20020192678A1 1138151.2
US-10-071-766-10

Query Match 14.4%; Score 350.4; DB 9; Length 1326;
Best Local Similarity 99.7%; Pred. No. 5; Se-91;
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
OY 2053 AGCAATTCCTTAACCTGCTTCTTAATGAGATGCTTGGCAGCAGGCTCCTCAGCTGTGT 2112
DB 1326 AGCAATTCCTTAACCTGCTTCTTAATGAGATGCTTGGCAGCAGGCTCCTCAGCTGTGT 1267
OY 2113 GTACACGACGAGACAGTATCAGTACAGCCATACAGCTGTCCACACTGAAAGACGTG 2172
DB 1266 GTACACGACGAGACAGTATCAGTACAGCCATACAGCTGTCCACACTGAAAGACGTG 1207
OY 2173 TCTTAAACACCTGGAATCAAAATGCTTAATGATTAATAATCCAGACTACTCTCAG 2232
DB 1206 TCTTAAACACCTGGAATCAAAATGCTTAATGATTAATAATCCAGACTACTCTCAG 1147
OY 2233 CCTTAAAGCCTTTTATTCATTAATAAATGTAAGAGTGAATGAACTGGAACAATTT 2292
DB 1146 CCTTAAAGCCTTTTATTCATTAATAAATGTAAGAGTGAATGAACTGGAACAATTT 1087
OY 2293 AACTCAGACTGTGATTCAGAGTGGGAAACCTTATCTATCTGATTCAGACAGCA 2352
DB 1086 AACTCAGACTGTGATTCAGAGTGGGAAACCTTATCTATCTGATTCAGACAGCA 1027
OY 2353 CACTTATGATTCCTCCCAAACTAATGATTAATAATAATAATAATAATAATAATAATA 2404
DB 1026 CACTTATGATTCCTCCCAAACTAATGATTAATAATAATAATAATAATAATAATAATA 975
```

RESULT 3
US-10-198-846-11070
Sequence 11070, Application US/10198846
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198, 846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11070
LENGTH: 1292
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1287, 1288, 1289, 1290, 1291, 1292
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070

Query Match 7.8%; Score 189.2; DB 9; Length 1292;
Best Local Similarity 62.4%; Pred. No. 3.2e-44;
Matches 296; Conservative 0; Mismatches 178; Indels 0; Gaps 0;
QY 349 TGGGCTTCTAGCTGAGCGTGTGCTGACCCGCTGTGAGACAGTACGAGAACCTGCGCT 408
DB 318 TAGGGTTTATCTACTCTGAGTGAACCGATGCTGAGACAGTGTGATTTGCTT 377
QY 409 GGGCCGACCGCTCTAGACCTGCTGGGCTTGTGAGAGGCAAGCAGCAGCAGCC 468
DB 378 GGGCCGACGAGCTTAATGCTCTAGCAGTGTTCAGGAGCGAGCAGCAGCAGG 437
QY 469 GGGCTGCTGGGCGCAGCTATCCCTAGCCAGCCAGCTGGGCAAGCTGCTATCTCGCA 528
DB 438 GGGCTGCTGAGAGAGAGCTGATGCGGTAGCTCACTCTGCTGCTATCTTCTGCT 497
QY 529 GGGTCAAGCAGCAGCTCTCAAGCGCTTCCCAAGGCGCCAGCAGCTGCTGAGAGGCT 588
DB 498 CGGTGAGCACTGCTGTGTACAAAGATTTCCCAAGTGAAGCAGCGTGTGAGAGGTT 557
QY 589 TTTATGCTCGGCGAGAACAGCAGTGTGAGAACTGAGCCTACCAACCAACATGTTCT 648
DB 558 TTTATGCAACAGATGAAAGAAATTTATCAACCACTCAAGTCTCTCATGCAATATT 617
QY 649 GGGTGGCTGGGTGTGTGCTTGCACCTGTCAATGAAGCGGTGCTTGAAGCTGCAATCC 708
DB 618 GGGTTCATCATCTGCTGTTGAAATCTTGCAACTGAAGCCGGAATGAAGTGAATCA 677
QY 709 GGGACCTATCTGCTGCTGAGAGCTGCTGAGAGAGATGAACACCTGCGTACAGTGTG 768
DB 678 GAGACAGTGTGATGCTGCAATCAATGATGATGATGATGATGATGATGATGATGATG 737
QY 769 GACACCTGATGCTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATG 822
DB 738 GCCTTTATTCGGTTATGACTGGGTGGATTCGGTGGTTACACCCAGTAG 791

RESULT 4
US-09-880-107-2174/c
Sequence 2174, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darcia T.
APPLICANT: Vockley, Joseph G.

APPLICANT: Scherf, Iwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-MO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2174
LENGTH: 1198
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 L20941
US-09-880-107-2174

Query Match 7.2%; Score 175; DB 10; Length 1198;
Best Local Similarity 100.0%; Pred. No. 4e-40;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2230 CAGCCTTATGCTTTTATCATTAATAAAGTGAAGCTAGACAGTGAACCTTGAACA 2289
DB 1198 CAGCCTTATGCTTTTATCATTAATAAAGTGAAGCTAGACAGTGAACCTTGAACA 1139
QY 2290 TTTATGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2349
DB 1138 TTTATGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1079
QY 2350 CCACACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2404
DB 1078 CCACACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1024

RESULT 5
US-09-768-826-16
Sequence 16, Application US/09768826
Patent No. US20020012966A1
GENERAL INFORMATION:
APPLICANT: Shi et al.
TITLE OF INVENTION: 18 human secreted proteins
FILE REFERENCE: PFS12P1
CURRENT APPLICATION NUMBER: US/09/768,826
CURRENT FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: PCT/US00/22350
PRIOR FILING DATE: 2000-08-15
PRIOR APPLICATION NUMBER: 60/148,759
PRIOR FILING DATE: 1999-08-16
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 1350
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1135)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1148)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1166)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1174)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1181)
OTHER INFORMATION: n equals a,t,g, or c


```
NAME/KEY: SITE
LOCATION: (1209)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1229)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1266)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1285)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1287)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1290)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1295)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1305)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1341)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1343)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1345)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1348)
OTHER INFORMATION: n equals a,t,g, or c
US-09-768-826-16

Query Match 5.2%; Score 126.8; DB 10; Length 1350;
Best Local Similarity 69.9%; Pred. No. 4.1e-26;
Matches 186; Conservative 0; Mismatches 77; Indels 3; Gaps 1;

OY 1174 GGTGGAGAGCAGCATCAACCCCTTTGGAGAGATGATGATTTGGAGACCACTG 1233
DB 436 GGTGGCTGACAGATCAATCAACCCATTTGGAGATGATGAGACTTTGAGCAATCA 495
OY 1234 GATTGTGACAGGAATTTGACAGTGTCCCTGTGGCTGTGATGAGATGACACGAGCCT 1293
DB 496 GCTCTAGACCGCACTGACAGTGTCCCTGTATCCGAGCAATATACAGAACT 555
OY 1294 GCTTGGATGAGCCGAGCATGATGAAATAGCCGAGCCACACCCCTACAC--- 1350
DB 556 TCCCTCCCTGAGAGGACCACTGATGAGAGCAGCCGACCCACCTACACTGT 615
OY 1351 AGCTGCTTCCGCGCTTCCGCTGAGCCCTTTATGAGCTCCACCTTCAACATGAGCCT 1410
DB 616 GGCACGCGGCGCGAGCTCTCTGCGGCCCATTCCTTGCGCTCACCTTCAACCTGCGCAT 675
OY 1411 GAACAAAGAGAGATGAGTTCCAGC 1436
DB 676 GAGCGACGACCCCTGAGCAGAGCCTGC 701

RESULT 6
US-10-198-846-11346
; Sequence 11346, Application US/10198846
; Publication No. US20030099974A1
```

```
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steimann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
PRIOR FILING DATE: 2002-07-18
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11346
LENGTH: 853
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 23, 27, 36, 37, 40, 41,
LOCATION: 43, 50, 51, 54, 57, 853
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11346

Query Match 4.8%; Score 116; DB 9; Length 853;
Best Local Similarity 68.2%; Pred. No. 4.3e-23;
Matches 161; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

OY 350 GGGCTCTAGCAGCAGCTGCTGCTGACCCGCTGAGACAGTACAGAGACCTGCGCTG 409
DB 272 GGGGTTTATGTTACTTCTGTTAGTGAACCATGTTGTAATTTGGCTG 331
OY 410 GCCCGACCGCTCATGAGCTGTGCTGCGGCTTCTGAGAGCAAGAGCAAGAGCCG 469
DB 332 GCCAGACAGCTATGTTCTCTCATCTACAGAGTTCACGAGAGCAGCAGCGGCG 391
OY 470 GCTGCTGCGGCGGACGCTCATCCGCTACGCCCACTGGGACAGTGTCTATCCGCGCAG 529
DB 392 CCTGCTTAGAAGAGCCTGATGCTACGTCATCTCCTGCTCATCTTTGCTC 451
OY 530 CGTACAGCCGAGCTACAGAGCCTTCCGACGCGCCACGACCTGCTGCAAGCAG 585
DB 452 GGTGACATCTGCTGTACAAAGATTTCCACAAATGACACGAGTGTGAAGAG 507

RESULT 7
US-10-198-846-2561/C
; Sequence 2561, Application US/10198846
; Publication No. US20030099974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steimann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
PRIOR FILING DATE: 2002-07-18
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2561
LENGTH: 462
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 390, 462
```

OTHER INFORMATION: n = A,T,C or G
US-10-198-846-2561

Query Match 4.3%; Score 104; DB 9; Length 462;
Best Local Similarity 67.3%; Pred. No. 9e-20;
Matches 146; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 349 TGGGCTTCTACGTGACGCTGTCGTGACCCGCTGTGGGAACCACTACGAGAACCTGCCGT 408
DB 255 TAGGGTTTATGTACTCTGTAGTGAACCGATGTGGAAACGATTGTGAAATTTGCCCT 196
QY 409 GGGCCGACCGCTCATACGCTGTGTCGGGCTTCGTGGAAGCAAGGACGACGAGGCC 468
DB 195 GGGCAGACAGCTAATGTCTCTCATCTACGACGCTGTCAAGGAAGCAGCAGCAGGCG 136
QY 469 GGGCTGCGGGCGGACGCTATCCGCTACGCAACCTGGGCAAGCTGCTCATCTGCCGA 528
DB 135 GCGTGTCTTAGAAGACGCTGATGGCTACGTAACGTAATCTCACCCTCCCTGCTCATCTTCCCT 76
QY 529 GCGTCACGACCGCAGTCTACAAAGCGCTTCCCGACG 564
DB 75 CGGTGACACACTGCTGTACTCGCGCGCCACCGG 40

RESULT 8
US-10-198-846-7843
Sequence 7843, Application US/10198846
Publication No. US20030099974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198, 846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7843
LENGTH: 615
TYPE: DNA
ORGANISM: Homo sapiens
US-10-198-846-7843

Query Match 4.2%; Score 102; DB 9; Length 615;
Best Local Similarity 67.3%; Pred. No. 4.1e-19;
Matches 144; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 349 TGGGCTTCTACGTGACGCTGTCGTGACCCGCTGTGGGAACCACTACGAGAACCTGCCGT 408
DB 353 TAGGGTTTATGTACTCTGTAGTGAACCGATGTGGAAACGATTGTGAAATTTGCCCT 412
QY 409 GGGCCGACCGCTCATACGCTGTGTCGGGCTTCGTGGAAGCAAGGACGACGAGGCC 468
DB 413 GGGCAGACAGCTAATGTCTCTCATCTACGACGCTGTCAAGGAAGCAGCAGCAGGCG 472
QY 469 GGGCTGCGGGCGGACGCTATCCGCTACGCAACCTGGGCAAGCTGCTCATCTGCCGA 528
DB 473 GCGTGTCTTAGAAGACGCTGATGGCTACGTAACGTAATCTCACCCTCCCTGCTCATCTTCCCT 532
QY 529 GCGTCACGACCGCAGTCTACAAAGCGCTTCCCGACG 562
DB 533 CGGTGACACACTGCTGTACTCGCGCGCCACCGG 566

RESULT 9
US-09-764-872-195
Sequence 195, Application US/09764872

Publication No. US20030050231A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA125
CURRENT APPLICATION NUMBER: US/09/764, 872
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 957
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 195
LENGTH: 507
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (490)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-872-195

Query Match 4.0%; Score 98; DB 9; Length 507;
Best Local Similarity 62.0%; Pred. No. 5.2e-18;
Matches 155; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 103 CCATGACCATCTTACACAAAGGCGCTAATGCGCTTAGGCTCTTCCGCC 162
DB 35 CCATGACGCTTCTATACACTCTCAAGTGGCGAGGCCGCTTGGAGTTCTCTGCGC 94
QY 163 TGTCTGTGTGCTGGGCGGACATCTACAGCTGTATATGCGAGTTCTTAACTTCC 222
DB 95 TGTCTTCCGCTGAGGGAAGCANTCTACAAAGCTCTCTACAAAGAAATTCCTCTTGG 154
QY 223 TGTCTGTCTATCATATCCGCTTATTTATAGGCTGCGCTCAGGAAGAACAGC 282
DB 155 GGGCTGTGTAGCTGTGTGTGATGATCACTACCGCTGCTGTGACCCAGAGAGAGT 214
QY 283 TGATTTGAGAAAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 342
DB 215 ACGTATATCTCAGGTGGCGCGGATCTGACCAACGCTCAGCAGACCTCATCTCTGTCT 274
QY 343 TCGTCTGCGG 352
DB 275 TTGTATTGGG 284

RESULT 10
US-09-263-959-1/C
Sequence 1, Application US/09263959
Patent No. US20020150891A1
GENERAL INFORMATION:
APPLICANT: Hood, Leroy E.
APPLICANT: Rowen, Lee
APPLICANT: Koop, Ben F.
TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH
NUMBER OF SEQUENCES: 1279
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/263,959
FILING DATE: 05-MAR-1999
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.

```

; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 920010.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 684973 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-263-959-1

Query Match
Best Local Similarity 3.7%; Score 89; DB 9; Length 31994;
Matches 98; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

1079 AGAAGACCCAGGCGTGGTGGCGACACCTGTATCCAGTACTCGGAGGCTGAGGC 1138
118384 AAAATATCAGCCAGGCGTGGTGGCGACACCTGTATCCAGTACTCGGAGGCTGAGGC 118325
1139 AGAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCAACCC 1198
118324 AGAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCAACCC 118265
1139 TT 1200
118264 TT 118263

RESULT 11
US-09-764-904-71/c
; Sequence 71, Application US/09764904
; Patent No. US20020173454A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122
; CURRENT APPLICATION NUMBER: US/09/764,904
; PRIOR APPLICATION DATA: 2001-01-17
; PRIOR APPLICATION DATA: removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 71
; LENGTH: 31994
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-904-71

Query Match
Best Local Similarity 3.7%; Score 89; DB 9; Length 31994;
Matches 98; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

1081 AAAGACCCAGGCGTGGTGGCGACACCTGTATCCAGTACTCGGAGGCTGAGGCAG 1140
6484 AAATATGACGAGGCGTGGTGGCGACACCTGTATCCAGTACTCGGAGGCTGAGGCAG 6425
1141 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 1193
6424 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 6372

RESULT 12
US-10-091-548-71/c
; Sequence 71, Application US/10091548
; Publication No. US20030049703A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122C1
; CURRENT APPLICATION NUMBER: US/10/091,548
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 137
```

```

; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 71
; LENGTH: 31994
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-091-548-71

Query Match
Best Local Similarity 3.7%; Score 89; DB 9; Length 31994;
Matches 98; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

1081 AAAGACCCAGGCGTGGTGGCGACACCTGTATCCAGTACTCGGAGGCTGAGGCAG 1140
6484 AAATATGACGAGGCGTGGTGGCGACACCTGTATCCAGTACTCGGAGGCTGAGGCAG 6425
1141 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 1193
6424 GAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGGCAGAGCAGCTCATCA 6372

RESULT 13
US-10-074-095-599/c
; Sequence 599, Application US/10074095
; Publication No. US20030077704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: 09/764,860
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
```

PRIOR APPLICATION NUMBER: 60/224,518
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/236,369
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/224,519
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/220,964
 PRIOR FILING DATE: 2000-07-26
 PRIOR APPLICATION NUMBER: 60/241,809
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/249,299
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/236,327
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/241,785
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/244,617
 PRIOR FILING DATE: 2000-11-01
 PRIOR APPLICATION NUMBER: 60/225,268
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/236,368
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/251,856
 PRIOR FILING DATE: 2000-12-08
 PRIOR APPLICATION NUMBER: 60/251,868
 PRIOR FILING DATE: 2000-12-08
 PRIOR APPLICATION NUMBER: 60/229,344
 PRIOR FILING DATE: 2000-09-01
 PRIOR APPLICATION NUMBER: 60/234,997
 PRIOR FILING DATE: 2000-09-25
 PRIOR APPLICATION NUMBER: 60/229,343
 PRIOR FILING DATE: 2000-09-01
 PRIOR APPLICATION NUMBER: 60/229,345
 PRIOR FILING DATE: 2000-09-01
 PRIOR APPLICATION NUMBER: 60/229,287
 PRIOR FILING DATE: 2000-09-01
 PRIOR APPLICATION NUMBER: 60/229,513
 PRIOR FILING DATE: 2000-09-05
 PRIOR APPLICATION NUMBER: 60/231,413
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/229,509
 PRIOR FILING DATE: 2000-09-05
 PRIOR APPLICATION NUMBER: 60/236,367
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/237,039
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/237,038
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/236,370
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/236,802
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/237,037
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/237,040
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/240,960
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/239,935
 PRIOR FILING DATE: 2000-10-13
 PRIOR APPLICATION NUMBER: 60/239,937
 PRIOR FILING DATE: 2000-10-13
 PRIOR APPLICATION NUMBER: 60/241,787
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/246,474
 PRIOR FILING DATE: 2000-11-08
 PRIOR APPLICATION NUMBER: 60/246,532
 PRIOR FILING DATE: 2000-11-08
 PRIOR APPLICATION NUMBER: 60/249,216
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,210
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/226,681

PRIOR FILING DATE: 2000-08-22
 PRIOR APPLICATION NUMBER: 60/225,759
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/225,213
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/227,182
 PRIOR FILING DATE: 2000-08-22
 PRIOR APPLICATION NUMBER: 60/225,214
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/235,836
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: 60/230,438
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/215,135
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: 60/225,266
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/249,218
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,208
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,213
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,212
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,207
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,245
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,244
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,217
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,211
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,215
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,264
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,214
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,297
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/232,400
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/231,242
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/232,081
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/232,080
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/231,414
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/231,244
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 60/233,064
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/233,063
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/232,397
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/232,399
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/232,401
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: 60/241,808
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/241,826
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/241,786
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/241,221
 PRIOR FILING DATE: 2000-10-20

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 09:29:22 ; Search time 488.416 Seconds

(without alignments)
11199.691 Million cell updates/sec

Title: US-09-622-964-4
Perfect score: 2429
Sequence: 1 caggagagtcacacagccta.....aaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|----------|--------------------|
| 1 | 2429 | 100.0 | 2429 | AAZ21228 | Human CGICE long f |
| 2 | 2013 | 82.9 | 2229 | AAZ21227 | Human CGICE short |
| 3 | 1244.2 | 51.2 | 1263 | AAV99722 | Human adult retina |
| 4 | 658.2 | 27.1 | 1916 | AAZ21229 | Mouse CGICE cDNA s |
| 5 | 642.2 | 26.4 | 16125 | AAZ21226 | Human CGICE genom |
| 6 | 642.2 | 26.4 | 16650 | ABAI4559 | Human nervous syst |
| 7 | 642.2 | 26.4 | 18530 | ABAI4557 | Human nervous syst |
| 8 | 640.6 | 26.4 | 18537 | ABAI4558 | Human nervous syst |
| 9 | 587.8 | 24.2 | 1717 | ABL89697 | Human polynucleoti |

| | | | | | | |
|----|-------|------|--------|----|----------|---------------------|
| 10 | 476.4 | 19.6 | 7108 | 22 | ABAI4556 | Human nervous syst |
| 11 | 294.6 | 12.1 | 305 | 22 | ABAI1094 | Human nervous syst |
| 12 | 237 | 9.8 | 2861 | 23 | ABLI0793 | Drosophila melanog |
| 13 | 234.8 | 9.7 | 424 | 20 | AAV86976 | EST clone BK112. |
| 14 | 193 | 7.9 | 1608 | 23 | ABLI2609 | Drosophila melanog |
| 15 | 188.2 | 7.7 | 10760 | 23 | ABLI0792 | Drosophila melanog |
| 16 | 175 | 7.2 | 1198 | 24 | ABK83853 | Human cDNA differe |
| 17 | 175 | 7.2 | 1198 | 24 | ABN95676 | Gene #2174 used to |
| 18 | 175 | 7.2 | 1234 | 23 | AA567682 | DNA encoding novel |
| 19 | 175 | 7.2 | 1279 | 23 | AA585665 | DNA encoding novel |
| 20 | 175 | 7.2 | 1526 | 23 | AA591587 | DNA encoding novel |
| 21 | 154.8 | 6.4 | 1526 | 23 | AA572796 | DNA encoding novel |
| 22 | 150.4 | 6.2 | 3592 | 23 | ABLI2636 | Drosophila melanog |
| 23 | 146.8 | 6.0 | 1345 | 23 | ABLI2637 | Drosophila melanog |
| 24 | 126.8 | 5.2 | 1350 | 22 | AA76848 | Human secreted pro |
| 25 | 120.4 | 5.0 | 2196 | 23 | ABLI6195 | Drosophila melanog |
| 26 | 115 | 4.7 | 2104 | 23 | AA572795 | DNA encoding novel |
| 27 | 104.8 | 4.3 | 3785 | 23 | ABLI2608 | Drosophila melanog |
| 28 | 100.8 | 4.1 | 392 | 22 | AAI11210 | Human breast cance |
| 29 | 100.8 | 4.1 | 401 | 22 | AAI20109 | Human breast cance |
| 30 | 99.8 | 4.1 | 527 | 22 | AAI17589 | Human breast cance |
| 31 | 98 | 4.0 | 507 | 22 | AA539532 | Human encoding nove |
| 32 | 98 | 4.0 | 507 | 22 | AAK88431 | Human encoding nove |
| 33 | 90 | 3.7 | 267156 | 24 | ABL68560 | Human digestive sy |
| 34 | 89 | 3.7 | 514 | 22 | AAH13186 | Human cDNA clone (|
| 35 | 89 | 3.7 | 31994 | 22 | AA530619 | DNA encoding novel |
| 36 | 89 | 3.7 | 31994 | 22 | AA528165 | Genomic sequence # |
| 37 | 88.8 | 3.7 | 10366 | 22 | AAK71267 | Human immune/haema |
| 38 | 88.8 | 3.7 | 10366 | 22 | AAK90439 | Human digestive sy |
| 39 | 88.8 | 3.7 | 26427 | 22 | ABA20762 | Human nervous syst |
| 40 | 88.8 | 3.7 | 26427 | 22 | ABA20763 | Human nervous syst |
| 41 | 88.2 | 3.6 | 8056 | 22 | AAK79815 | Human immune/haema |
| 42 | 88 | 3.6 | 30568 | 22 | AAI37486 | Human musculoskele |
| 43 | 87.8 | 3.6 | 139904 | 24 | ABK83562 | Human cDNA differe |
| 44 | 87.6 | 3.6 | 19199 | 22 | AAK70995 | Human immune/haema |
| 45 | 87.4 | 3.6 | 370 | 14 | AAQ59541 | Human brain Expres |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | |
| ID | AAZ21228 standard; cDNA; 2429 BP. |
| AC | AAZ21228; |
| XX | |
| XX | |
| XX | |
| DT | 22-NOV-1999 (first entry) |
| DE | Human CGICE long form cDNA sequence. |
| XX | |
| XX | |
| KW | CGICE; Best's macular dystrophy; mutation; diagnosis; detection; |
| KW | BMD; age-related macular dystrophy; ss. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| FH | Key |
| FT | CDS |
| FT | Location/Qualifiers |
| FT | 105..1412 |
| FT | /*tag= a |
| XX | /product= "CGICE short form protein" |
| PN | W09943695-A1. |
| XX | |
| PD | 02-SEP-1999. |
| XX | |
| XX | |
| PF | 22-FEB-1999; 99WO-US03790. |
| XX | |
| PR | 25-FEB-1998; 98US-0075941. |
| PR | 18-DEC-1998; 98US-0112926. |
| XX | |
| PA | (MERI) MERCK & CO INC. |
| XX | (UYUP-) UNIV UPSALA. |

Petrushin K, Caskey CT, Metzker M, Wadelius C;
 WPI: 1999-540560/45.
 P-PSDB: AAY29954.

Human and mouse polynucleotides encoding CGICE polypeptides
 Claim 2; Fig 4; 67pp; English.

The present sequence represents the human CGICE cDNA sequence, which
 CC when mutated is responsible for Best's macular dystrophy (BMD).
 CC Polynucleotides encoding CGICE are useful for diagnosing whether a
 CC patient carries a mutation in the CGICE gene. Normal and mutated
 CC CGICE proteins are useful for identifying activators and/or inhibitors
 CC of these proteins. In order to treat BMD, the CGICE gene offers a
 CC simpler and cheaper method of diagnosing BMD without the need for the
 CC presence of the patient. The gene may also be useful to discovering
 CC the genetic cause of age-related macular dystrophy.

Sequence 2429 BP; 614 A; 694 C; 613 G; 508 T; 0 other;

Query Match 100.0%; Score 2429; DB 20; Length 2429;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAGGAGTCCACCAAGCTAGTCCAGACCTTCTGGGATCATGAGACCCACCTGGAA 60
 1 CAGGAGTCCACCAAGCTAGTCCAGACCTTCTGGGATCATGAGACCCACCTGGAA 60
 61 CCCACCTGACCCCAAGCCACCTGTCAGACCCACCTGTCGCTGCTGCTGCTGCTG 120
 61 CCCACCTGACCCCAAGCCACCTGTCAGACCCACCTGTCGCTGCTGCTGCTGCTG 120
 121 CAGGAGTCCACCAAGCTAGTCCAGACCTTCTGGGATCATGAGACCCACCTGGAA 180
 121 CAGGAGTCCACCAAGCTAGTCCAGACCTTCTGGGATCATGAGACCCACCTGGAA 180
 121 CAGGAGTCCACCAAGCTAGTCCAGACCTTCTGGGATCATGAGACCCACCTGGAA 180
 121 CAGGAGTCCACCAAGCTAGTCCAGACCTTCTGGGATCATGAGACCCACCTGGAA 180
 181 GCAGATCTCAACCTGATATGAGAGTCTTATCTTCTGCTGCTGCTGCTGCTGCTG 240
 181 GCAGATCTCAACCTGATATGAGAGTCTTATCTTCTGCTGCTGCTGCTGCTGCTG 240
 241 TCCGCTTATTTATAGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 241 TCCGCTTATTTATAGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 301 CTCTGATTTGGAAGCTAGTCCAGACCTTCTGGGATCATGAGACCCACCTGGAA 360
 301 CTCTGATTTGGAAGCTAGTCCAGACCTTCTGGGATCATGAGACCCACCTGGAA 360
 361 TACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 361 TACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 421 TATAGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 421 TATAGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 481 GCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 481 GCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 541 CAGTCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 541 CAGTCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 601 CAGACACAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 601 CAGACACAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 661 TGTGCTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 661 TGTGCTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 721 TGTGCTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 721 TGTGCTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780

721 TGTGCTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 721 TGTGCTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 781 CTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 781 CTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 841 GCTTCTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 841 GCTTCTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 901 GCTTCTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 901 GCTTCTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 961 GCTTCTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 961 GCTTCTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 1021 ATGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
 1021 ATGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
 1081 AAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 1081 AAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 1141 GAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 1141 GAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 1201 TGGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
 1201 TGGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
 1261 CCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1320
 1261 CCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1320
 1321 GATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
 1321 GATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
 1381 CTTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
 1381 CTTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
 1441 TCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 1441 TCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 1501 CAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
 1501 CAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
 1561 ATCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
 1561 ATCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
 1621 CAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
 1621 CAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
 1681 GTATCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
 1681 GTATCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
 1741 CTTCCCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
 1741 CTTCCCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
 1801 AGACAAAGCTTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
 1801 AGACAAAGCTTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860


```

Db      |||||
1801 AGACAAAGCTTAAAGCTGTGAGTCTGGGCGCAAGAAAGTTTGAAATGCTCTCACA 1860
QY      1861 GACGATGGGGCTTATGATGAGACCCAGAAATATCTCAATGAGAGGAAAACTGTGA 1920
Db      1861 GACGATGGGGCTTATGATGAGACCCAGAAATATCTCAATGAGAGGAAAACTGTGA 1920
QY      1921 GTTAACTGAGGATATGCGAGAGATCCCGGAAATCAGCTTAAGAACCTTGGACAA 1980
Db      1921 GTTAACTGAGGATATGCGAGAGATCCCGGAAATCAGCTTAAGAACCTTGGACAA 1980
QY      1981 ATCACCACCAACATATACACTACATCAAGATCAGATGATCTTATGGGCTTGA 2040
Db      1981 ATCACCACCAACATATACACTACATCAAGATCAGATGATCTTATGGGCTTGA 2040
QY      2041 AACAGGATGAAAGACATTCCTTAACCTGCTTCTTAATGGGAGTCTCCGACCAAGT 2100
Db      2041 AACAGGATGAAAGACATTCCTTAACCTGCTTCTTAATGGGAGTCTCCGACCAAGT 2100
QY      2101 CCTGACCTGTGTGACACAGCAGACAGATCATGATCAGCATACAGCTGTCCACA 2160
Db      2101 CCTGACCTGTGTGACACAGCAGACAGATCATGATCAGCATACAGCTGTCCACA 2160
QY      2161 CTGAAGAACTGTCTCTACACAGCCTGATCAATGATTAGCTTAATAGATAAATCCC 2220
Db      2161 CTGAAGAACTGTCTCTACACAGCCTGATCAATGATTAGCTTAATAGATAAATCCC 2220
QY      2221 AGACTACTTACAGCCTTATATGCTTTATTCATATAAACTGTGAAGCTACAGCAACA 2280
Db      2221 AGACTACTTACAGCCTTATATGCTTTATTCATATAAACTGTGAAGCTACAGCAACA 2280
QY      2281 TTGGAATATTAATGAGCTGATGATGAGAGCGGAAACCTTATGATCTGAT 2340
Db      2281 TTGGAATATTAATGAGCTGATGATGAGAGCGGAAACCTTATGATCTGAT 2340
QY      2341 CCAGACAGCCACACCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
Db      2341 CCAGACAGCCACACCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
QY      2401 CGTTAAAAAAGGAAAAAAAAAAAAA 2429
Db      2401 CGTTAAAAAAGGAAAAAAAAAAAAA 2429

RESULT 2
AAZ21227
ID      AAZ21227 standard; cDNA; 2229 bp.
XX      AC      AAZ21227;
XX      DT      22-NOV-1999 (first entry)
XX      DE      Human CGICE short form cDNA sequence.
XX      KW      CGICE; Best's macular dystrophy; mutation; diagnosis; detection;
XX      KM      BMD; age-related macular dystrophy; ss.
XX      OS      Homo sapiens.
FH      Key      Location/Qualifiers
FT      CDS      105..1862
FT      FT      /*tag- a
FT      FT      /product- "CGICE long form protein"
FT      FT      /transl_except- (pos:465..467,aa:Ser)
XX      PN      MO9943695-A1.
XX      PD      02-SEP-1999.
XX      PF      22-FEB-1999; 99WO-US03790.
XX      PR      25-FEB-1998; 98US-0075941.
XX      PR      18-DEC-1998; 98US-0112926.

```

```

XX      (MERI ) MERCK & CO INC.
PA      (UYUP-) UNIT UPSALA.
XX      Petukhin K, Caskey CT, Metzker M, Madellus C;
XX      WPI; 1999-540560/45.
DR      P-PSDB; AAY29953.
XX      Human and mouse polynucleotides encoding CGICE polypeptides
PS      Claim 2; Fig 2; 67pp; English.
XX      The present sequence represents the human CGICE cDNA sequence, which
XX      is responsible for Best's macular dystrophy (BMD).
XX      polynucleotides encoding CGICE are useful for diagnosing whether a
XX      patient carries a mutation in the CGICE gene. Normal and mutated
XX      CGICE proteins are useful for identifying activators and/or inhibitors
XX      of these proteins. In order to treat BMD, the CGICE gene offers a
XX      simpler and cheaper method of diagnosing BMD without the need for the
XX      presence of the patient. The gene may also be useful to discovering
XX      the genetic cause of age-related macular dystrophy.
SQ      Sequence 2229 bp; 575 A; 646 C; 532 G; 476 T; 0 other;
Query Match      82.9%; Score 2013; DB 20; Length 2229;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 2226; Conservative 0; Mismatches 0; Indels 203; Gaps 1;
QY      1 CAGGAGATCCACACAGCTTATGCTGAGATGATGATGATGATGATGATGATGATGAT 60
Db      1 CAGGAGATCCACACAGCTTATGCTGAGATGATGATGATGATGATGATGATGATGAT 60
QY      61 CCCACCTGACCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG 120
Db      61 CCCACCTGACCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG 120
QY      121 CAGCAGAGTGGTAAATGCGGCTTATGCTGCTTCTCCGCTGCTGCTGCTGCTGCTGCTG 180
Db      121 CAGCAGAGTGGTAAATGCGGCTTATGCTGCTTCTCCGCTGCTGCTGCTGCTGCTGCTG 180
QY      181 GCAGCTCTACAGCTGCTATATGCGAGATTTCTTATCTTCTGCTGCTGCTGCTGCTGCT 240
Db      181 GCAGCTCTACAGCTGCTATATGCGAGATTTCTTATCTTCTGCTGCTGCTGCTGCTGCT 240
QY      241 TCCGCTTATTTATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db      241 TCCGCTTATTTATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY      301 CTCTGTATTTGAGACAGCTATCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 360
Db      301 CTCTGTATTTGAGACAGCTATCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 360
QY      361 TGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db      361 TGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY      421 TCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db      421 TCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY      481 GCAGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db      481 GCAGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY      541 CAGTCTACAGGCTTCCCGACGCGCCAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db      541 CAGTCTACAGGCTTCCCGACGCGCCAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY      601 CAGAACACAGAGCTTGGAGAACTGAGCTTACACACACACACACACACACACACACAC 660
Db      601 CAGAACACAGAGCTTGGAGAACTGAGCTTACACACACACACACACACACACACACAC 660

```

| | | | |
|----|------|---|------|
| Oy | 661 | TGTGTTTGCACAACTGTCAATGAAGGCGTGCCTTGGAGGTGAAATCCGGGACCTAATCC | 720 |
| Db | 661 | TGTGTTTGCACAACTGTCAATGAAGGCGTGCCTTGGAGGTGAAATCCGGGACCTAATCC | 720 |
| Oy | 721 | TGCTCCAGAGCCTGCTGAGACGAGATGAACACTGTGGTACTGAGTGTGGACACTGTATG | 780 |
| Db | 721 | TGCTCCAGAGCCTGCTGAGACGAGATGAACACTGTGGTACTGAGTGTGGACACTGTATG | 780 |
| Oy | 781 | CCTACGACGTGATATGATATCCACGTGTATACACAGTGTGACTGTGGCGGTGTACA | 840 |
| Db | 781 | CCTACGACGTGATATGATATCCACGTGTATACACAGTGTGAGTGTGGCGGTGTACA | 840 |
| Oy | 841 | GCTTCTTCTGACTTGTCTAGTGGGCGGAGTTTGTGAACCCAGGCCACTACCTTG | 900 |
| Db | 841 | GCTTCTTCTGACTTGTCTAGTGGGCGGAGTTTGTGAACCCAGGCCACTACCTTG | 900 |
| Oy | 901 | GCCATATAGCTGAGACCTCGTGTGCGCGGTCTTCAAGTTCAGAGTCTTCTATGTTG | 960 |
| Db | 901 | GCCATATAGCTGAGACCTCGTGTGCGCGGTCTTCAAGTTCAGAGTCTTCTATGTTG | 960 |
| Oy | 961 | GCTGGCTGAAGGTGGGCTCTCCAGGGCCCTGCTGGGCTGAGGCAATGGCCAGAGGGGTC | 1020 |
| Db | 961 | GCTGGCTGAAGGTGGGCTCTCCAGGGCCCTGCTGGGCTGAGGCAATGGCCAGAGGGGTC | 1020 |
| Oy | 1021 | ATGGCCACAGACGTGCTTGAGACGAGATGATGTCAGAAAGAGAGTCTACAGGGTAG | 1080 |
| Db | 971 | ----- | 970 |
| Oy | 1081 | AAAGCAGCCAGGCGTGTGGCCACACCTGTATATCCACGTACATCGGAGGCTGAGCAG | 1140 |
| Db | 971 | ----- | 970 |
| Oy | 1141 | GAGAAATGCTTGAACCCGGAGGCGAGGTTGTGGTGAGAGCAGCTCATACCCCTT | 1200 |
| Db | 971 | -----GGTGGCAGAGCGCTCATACACCCCTT | 997 |
| Oy | 1201 | TGAGAGAGATGATGATTTTGGAGACCACTGGATTCGACAGAAATTTGCAGGTGC | 1260 |
| Db | 998 | TGGAGAGATGATGATGATTTTGGAGACCACTGGATTTGCGAGAAATTTGCAGGTGC | 1057 |
| Oy | 1261 | CCTATGGCTGTGATGATGACACACAGAGCTGCTCGATGGATGAGCCGGACATGTACTG | 1320 |
| Db | 1058 | CCTATGGCTGTGATGATGATGACACACAGAGCTGCTCGATGGATGAGCCGGACATGTACTG | 1117 |
| Oy | 1321 | GAATTAAGCCGAGCCACAGCCCCCTTACACAGCTGTTCCGCCAGTTCCGTCGAGCCGC | 1380 |
| Db | 1118 | GAATTAAGCCGAGCCGAGCCACAGCCCCCTTACACAGCTGTTCCGCCAGTTCCGTCGAGCCGC | 1177 |
| Oy | 1381 | CTTTATGGGCTCCACTTCACATCATGAGCTTGAACAAAGAGAGATGATTCAGGCCAA | 1440 |
| Db | 1178 | CTTTATGGGCTCCACTTCACATCATGAGCTTGAACAAAGAGAGATGATTCAGGCCAA | 1237 |
| Oy | 1441 | TCAGGAGCAGAGAGATGCTACGCTGGCATATYGGCCGCTTCTTAGGCTGTGAGTC | 1500 |
| Db | 1238 | TCAGGAGCAGAGAGATGCTACGCTGGCATATYGGCCGCTTCTTAGGCTGTGAGTC | 1297 |
| Oy | 1501 | CCATGATACCACTCTCCAGGCGCAATCAAGGACCAACTACTGTGGCCCAAGAGGA | 1560 |
| Db | 1298 | CCATGATACCACTCTCTCCAGGCGCAATCAAGGACCAACTACTGTGGCCCAAGAGGA | 1357 |
| Oy | 1561 | ATCCCTTCTCAGAGGCGCTGCCCCAAAAACCAAGAGGCCCAACAAACGTAAGTGGG | 1620 |
| Db | 1358 | ATCCCTTCTCAGAGGCGCTGCCCCAAAAACCAAGAGGCCCAACAAACGTAAGTGGG | 1417 |
| Oy | 1621 | CCAGGAGACACAAGGCCCTGAGAGCTTAAGGCTGAGAGCCCTCAAGTCTGGCCACT | 1680 |
| Db | 1418 | CCAGGAGAGACAAGGCCCTGAGAGCTTAAGGCTGAGAGCCCTCAAGTCTGGCCACT | 1477 |
| Oy | 1681 | GATTCAGAGCCAGGCTACTACAGTGGCCCAACAGAGCCGCTCAGCCCTCCATGTT | 1740 |
| Db | 1478 | GATTCAGAGCCAGGCTACTACAGTGGCCCAACAGAGCCGCTCAGCCCTCCATGTT | 1537 |
| Oy | 1741 | CTTCCCTCTGAACATACAGCCGCTTAAAGGCTTCAAGTGTCAAGGCAATGACACCAA | 1800 |

| | | | | |
|-----------------|--|--------------------------|--|------|
| Db | | 1538 | CTTCCCCCCTGGAACCACTACGCCCGCCTCAAAAGCTTCACACTGTGCACAGCATGACACCACA | 1597 |
| Oy | | 1801 | AGCAAAAAGTTTAAGAAGCTGTGAGTCTGTGGGGCCCAAAAAAGTTTGAATTCCTCTCAGA | 1860 |
| Db | | 1598 | AGACAAAAGCTTTAAGACACTGTGAGTCTGTGGGGCCCAAAAAAGTTTGAATTCCTCTCAGA | 1657 |
| Oy | | 1861 | GAGCGATGGGGCCTTGATGAGCACCCACCAAGATATCTCAAGTAGGAGGAAAACTGTGGA | 1920 |
| Db | | 1658 | GAGCGATGGGGCCTTGATGAGCACCCACCAAGATATCTCAAGTAGGAGGAAAACTGTGGA | 1717 |
| Oy | | 1921 | GTTTAACCTGACGGATATGCGCAGAGATCCCGAAAAATTCACCTCCMAAAGAACCTTTGGAGA | 1980 |
| Db | | 1718 | GTTTAACCTGACGGATATGCGCAGAGATCCCGAAAAATTCACCTCCMAAAGAACCTTTGGAGA | 1777 |
| Oy | | 1981 | ATCACCAACCAACATFACACACTACACTCAAAAGATACATGATGATCTTATTGGGGCCTTGGG | 2040 |
| Db | | 1778 | ATCACCAACCAACATFACACACTACACTCAAAAGATACATGATGATCTTATTGGGGCCTTGGG | 1837 |
| Oy | | 2041 | AAACAGGATTAACACATATTCCTAACCTGGCTTCCTAATGGGATGCTTCCGACCCAGGT | 2100 |
| Db | | 1838 | AAACAGGATTAACACATATTCCTAACCTGGCTTCCTAATGGGATGCTTCCGACCCAGGT | 1897 |
| Oy | | 2101 | CCTGACCTGTGTGTACACACAGCAGAGACACTGATCAGACGCATFACAGCTGTCCACA | 2160 |
| Db | | 1898 | CCTGACCTGTGTGTGTACACACAGCAGAGACACTGATCAGACGCATFACAGCTGTCCACA | 1957 |
| Oy | | 2161 | CTGAAGAAGCTGTCTCTACAAAGCCTGAATCAAATGGTTAGCTTAATGATAAAAAATCCC | 2220 |
| Db | | 1958 | CTGAAGAAGCAGTCTCTCTACAAAGCCTGAATCAAATGGTTAGCTTAATGATAAAAAATCCC | 2017 |
| Oy | | 2221 | AACACTACTTCAGCCTTTAATGCTTTTATTCATAAAAAATCTGTAAAGCTAATACCTGAACCA | 2280 |
| Db | | 2018 | AACACTACTTCAGCCTTTAATGCTTTTATTCATAAAAAATCTGTAAAGCTAATACCTGAACCA | 2077 |
| Oy | | 2281 | TTCGAAACATTTAATCTACAGACTCTGGATTCAGAGTCGGGAACCTTACTATCTGANT | 2340 |
| Db | | 2078 | TTCGAAACATTTAATCTACAGACTCTGGATTCAGAGTCGGGAACCTTACTATCTGANT | 2137 |
| Oy | | 2341 | CCAAGACAGCCACACACTTAGTACTGCCCCAAACTAATGAGTTTAAATACAAATACT | 2400 |
| Db | | 2138 | CCAAGACAGCCACACACTTAGTACTGCCCCAAACTAATGAGTTTAAATACAAATACT | 2197 |
| Oy | | 2401 | CGTTAAAAAATTT | 2429 |
| Db | | 2198 | CGTTAAAAAATTT | 2226 |
| RESULT 3 | | | | |
| ID | AAV9722 | standard; cDNA; 1263 bp. | | |
| XX | AAV9722; | | | |
| DT | 26-APR-1999 | (first entry) | | |
| XX | | | | |
| DE | Human adult retina secreted protein bk112_15 cDNA. | | | |
| KW | Secreted protein; human; retina; bk112_15; ds. | | | |
| OS | Homo sapiens. | | | |
| XX | | | | |
| FT | key | location/Qualifiers | | |
| FT | CDS | 100..885 | | |
| PN | | /tag= a | | |
| PD | MO3856909-A2. | | | |
| PF | 17-DEC-1998. | | | |
| XX | | | | |
| XX | 08-JUN-1998; | 98WO-US11822. | | |
| XX | 05-JUN-1998; | 98US-0092722. | | |

PR 11-JUN-1997; 97US-0873218.
 XX (GEM) GENETICS INST INC.
 XX Agostino MJ, Fechtel K, Howes SH, Jacobs K, Lavallie ER;
 PI McCoy JM, Racie LA, Spaulding V, Treacy M;
 XX WPI: 1999-080899/07.
 DR P-PSDB: AAW95345.
 XX New polynucleotides encoding secreted human proteins - derived from
 PT human foetal brain, adult testes, foetal kidney, adult thyroid or
 PT adult retina cDNA libraries
 PS Claim 12: Page 71-72: 113pp; English.
 XX This is the nucleotide sequence of cDNA clone bk12.15, which
 CC includes an open reading frame for a 261-amino acid polypeptide
 CC (see AAW95345). The clone was isolated from a human adult retina
 CC cDNA library using methods which are selective for cDNAs encoding
 CC secreted proteins, or was identified as encoding a secreted or
 CC transmembrane protein on the basis of computer analysis of the
 CC amino acid sequence of the encoding protein. Database searches
 CC indicate some sequence similarity to known sequences. The
 CC invention provides cDNA clones (see AAW9721-33) from human adult
 CC thyroid, adult retina, adult testis (see AAW95344-53). Each clone is
 CC individually available from deposit clone ATCC 98451 (see also
 CC AAW9734-43). The isolated polynucleotides (PNS) and proteins are
 CC predicted to have activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans
 CC and animals, although no supporting data is given. Suggested
 CC activities include nutritional, cytokine, cell proliferation or
 CC differentiation, immune stimulating (e.g. as vaccines) or immune
 CC suppressing, haematopoietic/regulating, tissue growth,
 CC activin/inhibin, chemotactic/chemokinetic, haemostatic,
 CC thrombolytic, receptor/ligand, antinflammatory, cadherin/tumour
 CC invasion suppressor, and tumour inhibition activities. The PNS are
 CC also stated to be useful for gene therapy.
 XX
 SO Sequence 1263 BP: 397 A; 342 C; 272 G; 252 T; 0 other;
 Query Match 51.2%; Score 1244.2; DB 20; Length 1263;
 Best Local Similarity 99.8%; Pred. No. 4.2e-280;
 Matches 1246; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1601 GCCAAACAGACCTTACGGGCGCCAGGAAGACACACAGCCCTGGAAGCTTAAGCTGTGGAC 1660
 DB 421 GCCAAACAGAACCTTAGGGGCGCCAGGAAGACACACAGCCCTGGAAGCTTAAGCTGTGGAC 480
 QY 1661 GCCTCAAGCTGCGCCACACTGATCAGAGCGCAGGCTACTACAGTCCGCCACAGAGGCC 1720
 DB 481 GCCTCAAGCTGCGCCACACTGATCAGAGCGCAGGCTACTACAGTCCGCCACAGAGGCC 540
 QY 1721 CTCAGGCCCTCCACTGTTCTTCCCTTGAACCATCAGCGCCCTCAAGCTTCACACT 1780
 DB 541 CTCAGGCCCTCCACTGTTCTTCCCTTGAACCATCAGCGCCCTCAAGCTTCACACT 600
 QY 1781 CTCAGGCCCTCCACTGTTCTTCCCTTGAACCATCAGCGCCCTCAAGCTTCACACT 1840
 DB 601 CTCAGGCCCTCCACTGTTCTTCCCTTGAACCATCAGCGCCCTCAAGCTTCACACT 660
 QY 1841 AGTTTGAATGCTCTCAGAGAGCGATGGGCGCTTGTATGAGCACCAGAAATATCTCAA 1900
 DB 661 AGTTTGAATGCTCTCAGAGAGCGATGGGCGCTTGTATGAGCACCAGAAATATCTCAA 720
 QY 1901 GTGAGAGGAAACCTTGGAACATCACAACCAACATACACTACCTCAAGATCAGATG 2020
 DB 721 GTGAGAGGAAACCTTGGAACATCACAACCAACATACACTACCTCAAGATCAGATG 780
 QY 1961 CTCAAAGAACCTTGGAAACATCACAACCAACATACACTACCTCAAGATCAGATG 2020
 DB 781 CTCAAAGAACCTTGGAAACATCACAACCAACATACACTACCTCAAGATCAGATG 840
 QY 2021 GATCCTTATTTGGGCGCTTGGAAACAGGAGATGACACATTTCTTAACCTGCTTCTAATGG 2080
 DB 841 GATCCTTATTTGGGCGCTTGGAAACAGGAGATGACACATTTCTTAACCTGCTTCTAATGG 900
 QY 2081 GGATGCTTGGGCGCGAGCGGCTCTCAGCTGTGTATGACACAGAGACACTGATCAGTCA 2140
 DB 901 GGATGCTTGGGCGCGAGCGGCTCTCAGCTGTGTATGACACAGAGACACTGATCAGTCA 960
 QY 2141 CAGCCATACAGCTGTCTCAGCTGTCTCAGCTGTCTCAGCTGTCTCAGCTGTCTCAG 2200
 DB 961 CAGCCATACAGCTGTCTCAGCTGTCTCAGCTGTCTCAGCTGTCTCAGCTGTCTCAG 1020
 QY 2201 GCTTATATGATTAATATCCAGACTACTTCAAGCTTAAAGCTTAAAGCTTAAAGCT 2260
 DB 1021 GCTTATATGATTAATATCCAGACTACTTCAAGCTTAAAGCTTAAAGCTTAAAGCT 1080
 QY 2261 GTGAAGCTAGACCTGATCAACATTTGAACATTTAATCAAGCTGTGATTCAGAGTGGGA 2320
 DB 1081 GTGAAGCTAGACCTGATCAACATTTGAACATTTAATCAAGCTGTGATTCAGAGTGGGA 1140
 QY 2321 ACCCTTAGTCTATCTGAATCCAGACAGCCACACCTTAGTACTGCTCCAAACTAATGA 2380
 DB 1141 ACCCTTAGTCTATCTGAATCCAGACAGCCACACCTTAGTACTGCTCCAAACTAATGA 1200
 QY 2381 GTTATATATATCAATATCTGTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAA 2429
 DB 1201 GTTATATATATCAATATCTGTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAA 1249

RESULT 4
 AA21229
 ID AA21229 standard; cDNA; 1916 BP.
 XX
 AC AA21229;
 XX
 DT 22-NOV-1999 (first entry)
 XX
 DE Mouse Ccl1c cDNA sequence.
 XX
 KW CG1cB: Best's macular dystrophy; mutation; diagnosis; detection;
 KM BMD; age-related macular dystrophy; ss.
 XX
 OS Mus sp.
 XX

| | | | | | | |
|----------|--|---------------------------|--|------|--|--|
| Oy | | 1644 | AGCTTAAGCGCTGTGGAGCCGCTTCAAGTCTGCCACACTGTATCAGAAGCCAGGCTACTACA | 1703 | | |
| Dd | | 1332 | AATTTAAAGGCTGTGGACTCTTGGAATGTCTTCAGAGTTAAGAGGAGGCTGCCATT | 1391 | | |
| Oy | | 1704 | GTGCCCCAGACGCCCCCTCAGCCCGACTCC | 1734 | | |
| Dd | | 1392 | GTGGCCACAGCACACCAGCAGCCACCTTAC | 1422 | | |
| RESULT 5 | | | | | | |
| ID | AA221226 | standard; DNA; 16125 BP. | | | | |
| XX | AA221226; | | | | | |
| XX | 22-NOV-1999 | (first entry) | | | | |
| DE | Human CG1CE genomic DNA sequence. | | | | | |
| XX | CG1CE; Best's macular dystrophy; mutation; diagnosis; detection; | | | | | |
| KW | BMD; age-related macular dystrophy; ss. | | | | | |
| OS | Homo sapiens. | | | | | |
| FH | Key | Location/Qualifiers | | | | |
| FT | Intron | 1..1513 | | | | |
| FT | | /tag= a | | | | |
| FT | | /number= 1 | | | | |
| FT | exon | 1514..1570 | | | | |
| FT | | /tag= b | | | | |
| FT | | /number= 1 | | | | |
| FT | Intron | 1571..3278 | | | | |
| FT | | /tag= c | | | | |
| FT | | /number= 2 | | | | |
| FT | exon | 3279..3466 | | | | |
| FT | | /tag= d | | | | |
| FT | | /number= 2 | | | | |
| FT | CDS | 3315..15787 | | | | |
| FT | | /tag= e | | | | |
| FT | | /product= "CG1CE protein" | | | | |
| FT | | /note= "contains introns" | | | | |
| FT | Intron | 3467..6616 | | | | |
| FT | | /tag= f | | | | |
| FT | | /number= 3 | | | | |
| FT | exon | 6617..6711 | | | | |
| FT | | /tag= g | | | | |
| FT | | /number= 3 | | | | |
| FT | Intron | 6712..7227 | | | | |
| FT | | /tag= h | | | | |
| FT | | /number= 4 | | | | |
| FT | exon | 7228..7461 | | | | |
| FT | | /tag= i | | | | |
| FT | | /number= 4 | | | | |
| FT | Intron | 7462..8353 | | | | |
| FT | | /tag= j | | | | |
| FT | | /number= 5 | | | | |
| FT | exon | 8354..8508 | | | | |
| FT | | /tag= k | | | | |
| FT | | /number= 5 | | | | |
| FT | Intron | 8509..8896 | | | | |
| FT | | /tag= l | | | | |
| FT | | /number= 6 | | | | |
| FT | exon | 8897..8974 | | | | |
| FT | | /tag= m | | | | |
| FT | | /number= 6 | | | | |
| FT | Intron | 8975..9676 | | | | |
| FT | | /tag= n | | | | |
| FT | | /number= 7 | | | | |
| FT | exon | 9677..10032 | | | | |
| FT | | /tag= o | | | | |
| FT | | /number= 7 | | | | |
| FT | Intron | 10033..11212 | | | | |
| FT | | /tag= p | | | | |

| Query Match | Best Local Similarity | 26.4%: Score 642.2; DB 20: Length 16125; | Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0; |
|--|-----------------------|--|--|
| 1405 CAGCCTAACAAGAGGATGAGATTCACAGCCCAATCAGAGAGAGGAGATGCTCA | 1464 | | |
| 13916 CAGCCTAACAAGAGGATGAGATTCACAGCCCAATCAGAGAGAGGAGATGCTCA | 13975 | | |
| 1465 CGCTGGCATCATTTGGCCGCTTCCTAGGCTCAGAGCCCATGATCATCAATCCCTCCAGGGCC | 1524 | | |
| 13976 CGCTGGCATCATTTGGCCGCTTCCTAGGCTCAGAGCCCATGATCATCAATCCCTCCAGGGCC | 14035 | | |
| 1525 AACTCAAGGACCAAACTACTGTGGCCCAAGAGGGAATCCCTTCTCCAGAGAGGCGCTGCC | 1584 | | |
| 14036 AACTCAAGGACCAAACTACTGTGGCCCAAGAGGGAATCCCTTCTCCAGAGAGGCGCTGCC | 14095 | | |
| 1585 CAAAACCCACAAGCAGCCCAACAGACGTTAAGGGCCCAAGAGACCAACAGCCCTGGGA | 1644 | | |

DB 14096 CAAAACCAAGGCGCCCAACAGACGTTAGGGCCAGGAGACAAAGGCTGGAA 14155
 QY 1645 GCTTAAGGCTGTGAGAGGCTTCAAGTCTGGCCACTGTATCAGAGGCCAGGCTACTACAG 1704
 DB 14156 GCTTAAGGCTGTGAGAGGCTTCAAGTCTGGCCACTGTATCAGAGGCCAGGCTACTACAG 14215
 QY 1705 TCGCCCAAGAGAGGCGCCCTCAGCCCACTCCATGTTCTTCCCTTAGAACCATCAGCGCC 1764
 DB 14216 TCGCCCAAGAGAGGCGCCCTCAGCCCACTCCATGTTCTTCCCTTAGAACCATCAGCGCC 14275
 QY 1765 GTCAAAAGCTTCAAGTGTCAAGGCTATGACACCAAGAACAAAGCTTAAGCTGTAG 1824
 DB 14276 GTCAAAAGCTTCAAGTGTCAAGGCTATGACACCAAGAACAAAGCTTAAGCTGTAG 14335
 QY 1825 TTCTGGGGCCAAAGATTTTGAATGCTCTCAGAGAGCGATGGGCGCTTGTAGAGCA 1884
 DB 14336 TTCTGGGGCCAAAGATTTTGAATGCTCTCAGAGAGCGATGGGCGCTTGTAGAGCA 14395
 QY 1885 CCCAGAGATATCTCAAGTGTAGAGAGAAAGTGTGAGTTTACCTGACGATATGCCAGA 1944
 DB 14396 CCCAGAGATATCTCAAGTGTAGAGAGAAAGTGTGAGTTTACCTGACGATATGCCAGA 14455
 QY 1945 GATCCCCGAAAAATCAGCTTCAAAACCTTGGAAACATCCCAACCATACACTAC 2004
 DB 14456 GATCCCCGAAAAATCAGCTTCAAAACCTTGGAAACATCCCAACCATACACTAC 14515
 QY 2005 ACTCAAGATCAGATGATCTTATTTGGGCTTGGAAAGAGGATG 2051
 DB 14516 ACTCAAGATCAGATGATCTTATTTGGGCTTGGAAAGAGGATG 14562

RESULT 6
 ABA14559 standard; DNA: 16650 BP.

ID ABA14559;
 AC ABA14559;
 DT 23-JAN-2002 (first entry)
 XX Human nervous system related polynucleotide seq ID NO 6890.
 DE Human: noctropic; neuroprotective; cytosolic; dermatological; virucide;
 KW immunosuppressive; antineoplastic; anti-HIV; antibacterial; vulnery;
 KW antiparkinsonian; antispasmodic; antianemic; antitubercular; cancer;
 KW antineoplastic; hepatotoxic; cerebroprotective; antineoplastic;
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 OS Homo sapiens.
 XX
 XX WO200159063-A2.
 PN 16-AUG-2001.
 PD 17-JAN-2001, 2001WO-US01334.
 PF 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205467.
 PR 07-JUN-2000; 2000US-0209451.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216847.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0227009.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234937.
 PR 25-SEP-2000; 2000US-0234938.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241185.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.

KM anti-parasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; neoplastic; gene therapy; vaccine; ds.

XX Homo sapiens.

PN WO200159063-A2.

PD 16-AUG-2001.

XX 17-JAN-2001: 2001MO-US01334.

XX 31-JAN-2000: 2000US-0179065.
PR 04-FEB-2000: 2000US-0180628.
PR 24-FEB-2000: 2000US-0184664.
PR 02-MAR-2000: 2000US-0186350.
PR 16-MAR-2000: 2000US-0189874.
PR 17-MAR-2000: 2000US-0190076.
PR 18-APR-2000: 2000US-0198123.
PR 19-MAY-2000: 2000US-0205515.
PR 07-JUN-2000: 2000US-0209467.
PR 28-JUN-2000: 2000US-0214886.
PR 30-JUN-2000: 2000US-0215135.
PR 07-JUL-2000: 2000US-0216647.
PR 07-JUL-2000: 2000US-0216880.
PR 11-JUL-2000: 2000US-0217487.
PR 11-JUL-2000: 2000US-0217496.
PR 14-JUL-2000: 2000US-0218290.
PR 26-JUL-2000: 2000US-0220963.
PR 26-JUL-2000: 2000US-0220964.
PR 14-AUG-2000: 2000US-0224518.
PR 14-AUG-2000: 2000US-0224519.
PR 14-AUG-2000: 2000US-0225213.
PR 14-AUG-2000: 2000US-0225214.
PR 14-AUG-2000: 2000US-0225266.
PR 14-AUG-2000: 2000US-0225267.
PR 14-AUG-2000: 2000US-0225268.
PR 14-AUG-2000: 2000US-0225270.
PR 14-AUG-2000: 2000US-0225447.
PR 14-AUG-2000: 2000US-0225757.
PR 14-AUG-2000: 2000US-0225758.
PR 14-AUG-2000: 2000US-0225759.
PR 18-AUG-2000: 2000US-0226279.
PR 22-AUG-2000: 2000US-0226681.
PR 22-AUG-2000: 2000US-0226688.
PR 22-AUG-2000: 2000US-0227009.
PR 23-AUG-2000: 2000US-0228924.
PR 30-AUG-2000: 2000US-0229287.
PR 01-SEP-2000: 2000US-0229343.
PR 01-SEP-2000: 2000US-0229344.
PR 01-SEP-2000: 2000US-0229345.
PR 05-SEP-2000: 2000US-0229509.
PR 05-SEP-2000: 2000US-0229513.
PR 06-SEP-2000: 2000US-0230437.
PR 06-SEP-2000: 2000US-0230438.
PR 08-SEP-2000: 2000US-0231142.
PR 08-SEP-2000: 2000US-0231243.
PR 08-SEP-2000: 2000US-0231244.
PR 08-SEP-2000: 2000US-0231413.
PR 08-SEP-2000: 2000US-0231414.
PR 08-SEP-2000: 2000US-0232080.
PR 08-SEP-2000: 2000US-0232081.
PR 12-SEP-2000: 2000US-0231968.
PR 14-SEP-2000: 2000US-0232397.
PR 14-SEP-2000: 2000US-0232398.
PR 14-SEP-2000: 2000US-0232399.
PR 14-SEP-2000: 2000US-0232400.
PR 14-SEP-2000: 2000US-0232401.
PR 14-SEP-2000: 2000US-0233063.
PR 14-SEP-2000: 2000US-0233064.
PR 14-SEP-2000: 2000US-0233065.
PR 21-SEP-2000: 2000US-0234223.
PR 21-SEP-2000: 2000US-0234274.

PR 25-SEP-2000: 2000US-0234997.
PR 25-SEP-2000: 2000US-0234998.
PR 26-SEP-2000: 2000US-0235484.
PR 27-SEP-2000: 2000US-0235834.
PR 27-SEP-2000: 2000US-0235836.
PR 29-SEP-2000: 2000US-0236327.
PR 29-SEP-2000: 2000US-0236367.
PR 29-SEP-2000: 2000US-0236368.
PR 29-SEP-2000: 2000US-0236369.
PR 29-SEP-2000: 2000US-0236370.
PR 02-OCT-2000: 2000US-0236802.
PR 02-OCT-2000: 2000US-0237037.
PR 02-OCT-2000: 2000US-0237038.
PR 02-OCT-2000: 2000US-0237039.
PR 02-OCT-2000: 2000US-0237040.
PR 02-OCT-2000: 2000US-0237041.
PR 13-OCT-2000: 2000US-0239933.
PR 13-OCT-2000: 2000US-0239937.
PR 20-OCT-2000: 2000US-0240960.
PR 20-OCT-2000: 2000US-0241785.
PR 20-OCT-2000: 2000US-0241786.
PR 20-OCT-2000: 2000US-0241787.
PR 20-OCT-2000: 2000US-0241808.
PR 20-OCT-2000: 2000US-0241809.
PR 20-OCT-2000: 2000US-0241826.
PR 20-OCT-2000: 2000US-0242221.
PR 20-OCT-2000: 2000US-0244617.
PR 01-NOV-2000: 2000US-0244617.
PR 08-NOV-2000: 2000US-0246474.
PR 08-NOV-2000: 2000US-0246475.
PR 08-NOV-2000: 2000US-0246476.
PR 08-NOV-2000: 2000US-0246477.
PR 08-NOV-2000: 2000US-0246478.
PR 08-NOV-2000: 2000US-0246523.
PR 08-NOV-2000: 2000US-0246524.
PR 08-NOV-2000: 2000US-0246525.
PR 08-NOV-2000: 2000US-0246526.
PR 08-NOV-2000: 2000US-0246527.
PR 08-NOV-2000: 2000US-0246528.
PR 08-NOV-2000: 2000US-0246532.
PR 08-NOV-2000: 2000US-0246609.
PR 08-NOV-2000: 2000US-0246610.
PR 08-NOV-2000: 2000US-0246611.
PR 08-NOV-2000: 2000US-0246613.
PR 17-NOV-2000: 2000US-0249207.
PR 17-NOV-2000: 2000US-0249208.
PR 17-NOV-2000: 2000US-0249209.
PR 17-NOV-2000: 2000US-0249210.
PR 17-NOV-2000: 2000US-0249211.
PR 17-NOV-2000: 2000US-0249212.
PR 17-NOV-2000: 2000US-0249213.
PR 17-NOV-2000: 2000US-0249214.
PR 17-NOV-2000: 2000US-0249215.
PR 17-NOV-2000: 2000US-0249216.
PR 17-NOV-2000: 2000US-0249217.
PR 17-NOV-2000: 2000US-0249218.
PR 17-NOV-2000: 2000US-0249244.
PR 17-NOV-2000: 2000US-0249245.
PR 17-NOV-2000: 2000US-0249246.
PR 17-NOV-2000: 2000US-0249265.
PR 17-NOV-2000: 2000US-0249297.
PR 17-NOV-2000: 2000US-0249299.
PR 17-NOV-2000: 2000US-0249300.
PR 17-NOV-2000: 2000US-0250391.
PR 01-DEC-2000: 2000US-0251160.
PR 01-DEC-2000: 2000US-0251160.
PR 05-DEC-2000: 2000US-0251988.
PR 05-DEC-2000: 2000US-0251988.
PR 06-DEC-2000: 2000US-0251479.
PR 06-DEC-2000: 2000US-0251479.
PR 08-DEC-2000: 2000US-0251865.
PR 08-DEC-2000: 2000US-0251869.
PR 08-DEC-2000: 2000US-0251869.
PR 08-DEC-2000: 2000US-0251869.
PR 11-DEC-2000: 2000US-0254097.

OY 1535 AACTCAAGGACCAACTACTGTGTGCCCCAAGAGGAAATCCCTTCACAGAGGGCGTCC 1584
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 12021 AACTCAAGGACCAACTACTGTGTGCCCCAAGAGGAAATCCCTTCACAGAGGGCGTCC 12080
 OY 1585 CAAAAACCAAGGACGCAACACAGAGGTTAGGGCCAGAGAACAAAGGCTGGAA 1644
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 12081 CAAAAACCAAGGACGCAACACAGAGGTTAGGGCCAGAGAACAAAGGCTGGAA 12140
 OY 1645 GCTTAGGCTGTGAGAGCCCTTCAAGTCTGCCCCACTGTATCAGAGGCCAGCTACTACAG 1704
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 12141 GCTTAGGCTGTGAGAGCCCTTCAAGTCTGCCCCACTGTATCAGAGGCCAGCTACTACAG 12200
 OY 1705 TGCCCCACAGAGCCCTTCAAGTCTGCCCCACTGTATCAGAGGCCAGCTACTACAG 1764
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 12201 TGCCCCACAGAGCCCTTCAAGTCTGCCCCACTGTATCAGAGGCCAGCTACTACAG 12260
 OY 1765 GTCAGAGCTTCAAGAGTGTACAGAGGATAGACACAAAGGCTTAAAGACTGTGAG 1824
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 12261 GTCAGAGCTTCAAGAGTGTACAGAGGATAGACACAAAGGCTTAAAGACTGTGAG 12320
 OY 1825 TTCTGGGGCCCAAGAAAGTTTAAATGCTCTCAGAGAGGATGGGCGCTGTATGAGCA 1884
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 12321 TTCTGGGGCCCAAGAAAGTTTAAATGCTCTCAGAGAGGATGGGCGCTGTATGAGCA 12380
 OY 1885 CCCAGAGATCTCAAGTGTAGAGAGAAACTGTGAGTTTAACTGAGGATATGCCAGA 1944
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 12381 CCCAGAGATCTCAAGTGTAGAGAGAAACTGTGAGTTTAACTGAGGATATGCCAGA 12440
 OY 1945 GATCCCCGAAATACCTTCAAAAGACCTTTGGAAACATACCAACCAATACACACTAC 2004
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 12441 GATCCCCGAAATACCTTCAAAAGACCTTTGGAAACATACCAACCAATACACACTAC 12500
 OY 2005 ACTCAAGATCATGATGATCTTATTTGGGCTTGGAAAGAGGATG 2051
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 12501 ACTCAAGATCATGATGATCTTATTTGGGCTTGGAAAGAGGATG 12547
 RESULT 9
 ABL89697 standard; cDNA: 1717 BP.
 ID ABL89697
 AC ABL89697;
 DT 24-MAY-2002 (first entry)
 DE Human polynucleotide SEQ ID NO 259.
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antifungal;
 KW vulnerability; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; gene; ss.
 XX
 OS Homo sapiens.
 PN WO200190304-A2.
 PD 29-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-US16450.
 PR 19-MAY-2000; 2000US-205515P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Birse CE, Rosen CA;
 PI
 XX WPI: 2002-122018/16.
 DR P-PSDB; ABB89288.
 XX
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative

PT disorders -
 XX
 PS Claim 4; SEQ ID NO 259; 2081bp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences.
 SO Sequence 1717 BP; 464 A; 460 C; 417 G; 367 T; 9 other:
 Query Match 24.2%; Score 587.8; DB 24; Length 1717;
 Best Local Similarity 98.8%; Pred. No. 5e-127;
 Matches 600; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
 OY 1445 GAGGAGAGAGAGATGCTACAGCTGCGATCATGTGCGCCCTTCTGAGGCTGAGTCCAT 1504
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1 GAGGAGAGAGAGATGCTACAGCTGCGATCATGTGCGCCCTTCTGAGGCTGAGTCCAT 60
 OY 1505 GATGACCAATCTCCCAAGGCAACACCAAGGACCAACTACTGTGCGCCCAAGAGGAAATCC 1564
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 61 GATGACCAATCTCCCAAGGCAACACCAAGGACCAACTACTGTGCGCCCAAGAGGAAATCC 120
 OY 1565 CTCTCTCAGAGAGGCTGCCCCAAACCAAGGACCAACCAAGGACCAACTACTGTGCGCCCAAG 1624
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 121 CTCTCTCAGAGAGGCTGCCCCAAACCAAGGACCAACCAAGGACCAACTACTGTGCGCCCAAG 179
 OY 1625 GAAGACAAACAGGCTGGAAGCTTAAAGCTGTGAGGACGCTTCAAGTCTGCGCCCAAGTAT 1684
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 180 GAAGACAAACAGGCTGGAAGCTTAAAGCTGTGAGGACGCTTCAAGTCTGCGCCCAAGTAT 239
 OY 1685 CAGAGGCAAGCTACTACAGTGTGCCCCACAGAGGCCCCCTCAGGCCCCCATGTCTTC 1744
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 240 CAGAGGCAAGCTACTACAGTGTGCCCCACAGAGGCCCCCTCAGGCCCCCATGTCTTC 299
 OY 1745 CCCCTAGAACCAATCAGGCGCGCTCAAAAGCTTCAAGTGTACAGAGGATAGACACCAAGAC 1804
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 300 CCCCTAGAACCAATCAGGCGCGCTCAAAAGCTTCAAGTGTACAGAGGATAGACACCAAGAC 359
 OY 1805 AAAAGCTTAAAGACTGTGAGTCTGAGGCGCCCAAGAAAGTTTGAATGCTCTCAGAGAGC 1864
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 360 AAAAGCTTAAAGACTGTGAGTCTGAGGCGCCCAAGAAAGTTTGAATGCTCTCAGAGAGC 419
 OY 1865 GATGGGCGCTTGTATGAGCAACCAAGATATCTCAAGTGTAGAGAGAAATGTGGAGTT 1924
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 420 GATGGGCGCTTGTATGAGCAACCAAGATATCTCAAGTGTAGAGAGAAATGTGGAGTT 479
 OY 1925 AACCTGAGGATATGAGCAACCAAGATATCTCAAGTGTAGAGAGAAATGTGGAGTT 1984
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 480 AACCTGAGGATATGAGCAACCAAGATATCTCAAGTGTAGAGAGAAATGTGGAGTT 539
 OY 1985 CCAACCAATACACACTACTACCTAAGATACATGTATCTTTTGGGCTTGGAAAC 2044
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 540 CCAACCAATACACACTACTACCTAAGATACATGTATCTTTTGGGCTTGGAAAC 599
 OY 2045 AGGAGTG 2051
 || ||
 DB 600 AGGCTG 606

RESULT 10
 ABA14556
 ID ABA14556 standard; DNA; 7108 BP.
 XX
 AC ABA14556;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 6887.
 XX
 KW Human; nootropic; neuroprotective; cytoskeletal; dermatological; virocidic;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; cancer;
 KW antiparkinsonian; antisticking; antianaemic; antirheumatic; antidiabetic;
 KW antihemagic; hepatoprotective; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antitumor;
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotoxic; gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200159063-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01334.
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198122.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 03-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236370.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 08-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.

| | | |
|----|-------------|-----------------|
| PR | 14-ANG-2000 | 2000US-022575.8 |
| PR | 14-ANG-2000 | 2000US-022575.8 |
| PR | 18-ANG-2000 | 2000US-022627.9 |
| PR | 18-ANG-2000 | 2000US-022627.9 |
| PR | 22-ANG-2000 | 2000US-022668.1 |
| PR | 22-ANG-2000 | 2000US-022668.1 |
| PR | 22-ANG-2000 | 2000US-022718.2 |
| PR | 22-ANG-2000 | 2000US-022718.2 |
| PR | 23-ANG-2000 | 2000US-022700.9 |
| PR | 23-ANG-2000 | 2000US-022700.9 |
| PR | 30-ANG-2000 | 2000US-022897.4 |
| PR | 30-ANG-2000 | 2000US-022897.4 |
| PR | 01-SEP-2000 | 2000US-022934.3 |
| PR | 01-SEP-2000 | 2000US-022934.3 |
| PR | 01-SEP-2000 | 2000US-022934.4 |
| PR | 01-SEP-2000 | 2000US-022934.5 |
| PR | 05-SEP-2000 | 2000US-022950.9 |
| PR | 05-SEP-2000 | 2000US-022951.3 |
| PR | 06-SEP-2000 | 2000US-023043.7 |
| PR | 06-SEP-2000 | 2000US-023043.8 |
| PR | 08-SEP-2000 | 2000US-023124.2 |
| PR | 08-SEP-2000 | 2000US-023124.4 |
| PR | 08-SEP-2000 | 2000US-023141.3 |
| PR | 08-SEP-2000 | 2000US-023141.4 |
| PR | 08-SEP-2000 | 2000US-023208.0 |
| PR | 08-SEP-2000 | 2000US-023208.1 |
| PR | 12-SEP-2000 | 2000US-023308.1 |
| PR | 12-SEP-2000 | 2000US-023308.6 |
| PR | 14-SEP-2000 | 2000US-023339.7 |
| PR | 14-SEP-2000 | 2000US-023339.8 |
| PR | 14-SEP-2000 | 2000US-023340.0 |
| PR | 14-SEP-2000 | 2000US-023340.1 |
| PR | 14-SEP-2000 | 2000US-023420.1 |
| PR | 14-SEP-2000 | 2000US-023306.3 |
| PR | 14-SEP-2000 | 2000US-023306.4 |
| PR | 14-SEP-2000 | 2000US-023306.5 |
| PR | 21-SEP-2000 | 2000US-023442.3 |
| PR | 21-SEP-2000 | 2000US-023442.3 |
| PR | 21-SEP-2000 | 2000US-023447.4 |
| PR | 21-SEP-2000 | 2000US-023449.7 |
| PR | 25-SEP-2000 | 2000US-023498.8 |
| PR | 25-SEP-2000 | 2000US-023535.0 |
| PR | 26-SEP-2000 | 2000US-023535.4 |
| PR | 26-SEP-2000 | 2000US-023535.4 |
| PR | 27-SEP-2000 | 2000US-023536.3 |
| PR | 27-SEP-2000 | 2000US-023536.7 |
| PR | 29-SEP-2000 | 2000US-023656.8 |
| PR | 29-SEP-2000 | 2000US-023656.8 |
| PR | 29-SEP-2000 | 2000US-023656.9 |
| PR | 29-SEP-2000 | 2000US-023657.0 |
| PR | 29-SEP-2000 | 2000US-023680.2 |
| PR | 02-OCT-2000 | 2000US-023703.7 |
| PR | 02-OCT-2000 | 2000US-023703.8 |
| PR | 02-OCT-2000 | 2000US-023703.9 |
| PR | 02-OCT-2000 | 2000US-023704.0 |
| PR | 13-OCT-2000 | 2000US-023993.5 |
| PR | 13-OCT-2000 | 2000US-023993.7 |
| PR | 13-OCT-2000 | 2000US-024095.0 |
| PR | 20-OCT-2000 | 2000US-024117.6 |
| PR | 20-OCT-2000 | 2000US-024117.6 |
| PR | 20-OCT-2000 | 2000US-024180.8 |
| PR | 20-OCT-2000 | 2000US-024180.9 |
| PR | 20-OCT-2000 | 2000US-024182.2 |
| PR | 20-OCT-2000 | 2000US-02422.1 |
| PR | 01-NOV-2000 | 2000US-024641.7 |
| PR | 01-NOV-2000 | 2000US-024641.7 |
| PR | 08-NOV-2000 | 2000US-024647.5 |
| PR | 08-NOV-2000 | 2000US-024647.6 |
| PR | 08-NOV-2000 | 2000US-024647.7 |
| PR | 08-NOV-2000 | 2000US-024647.8 |
| PR | 08-NOV-2000 | 2000US-024647.9 |
| PR | 08-NOV-2000 | 2000US-024652.5 |
| PR | 08-NOV-2000 | 2000US-024652.5 |
| PR | 08-NOV-2000 | 2000US-024652.6 |
| PR | 08-NOV-2000 | 2000US-024652.7 |
| PR | 08-NOV-2000 | 2000US-024652.8 |
| PR | 08-NOV-2000 | 2000US-024653.2 |
| PR | 08-NOV-2000 | 2000US-024653.2 |

| | | |
|----|-------------|----------------|
| PR | 08-NOV-2000 | 2000US-0246610 |
| PR | 08-NOV-2000 | 2000US-0246611 |
| PR | 08-NOV-2000 | 2000US-0246613 |
| PR | 17-NOV-2000 | 2000US-0249207 |
| PR | 17-NOV-2000 | 2000US-0249208 |
| PR | 17-NOV-2000 | 2000US-0249209 |
| PR | 17-NOV-2000 | 2000US-0249210 |
| PR | 17-NOV-2000 | 2000US-0249211 |
| PR | 17-NOV-2000 | 2000US-0249212 |
| PR | 17-NOV-2000 | 2000US-0249213 |
| PR | 17-NOV-2000 | 2000US-0249214 |
| PR | 17-NOV-2000 | 2000US-0249215 |
| PR | 17-NOV-2000 | 2000US-0249216 |
| PR | 17-NOV-2000 | 2000US-0249217 |
| PR | 17-NOV-2000 | 2000US-0249218 |
| PR | 17-NOV-2000 | 2000US-0249244 |
| PR | 17-NOV-2000 | 2000US-0249245 |
| PR | 17-NOV-2000 | 2000US-0249266 |
| PR | 17-NOV-2000 | 2000US-0249265 |
| PR | 17-NOV-2000 | 2000US-0249297 |
| PR | 17-NOV-2000 | 2000US-0249299 |
| PR | 17-NOV-2000 | 2000US-0249300 |
| PR | 01-DEC-2000 | 2000US-0250391 |
| PR | 01-DEC-2000 | 2000US-0251160 |
| PR | 05-DEC-2000 | 2000US-0251030 |
| PR | 05-DEC-2000 | 2000US-0251988 |
| PR | 05-DEC-2000 | 2000US-0256719 |
| PR | 06-DEC-2000 | 2000US-0251479 |
| PR | 08-DEC-2000 | 2000US-0251856 |
| PR | 08-DEC-2000 | 2000US-0251868 |
| PR | 08-DEC-2000 | 2000US-0251869 |
| PR | 08-DEC-2000 | 2000US-0251989 |
| PR | 08-DEC-2000 | 2000US-0251990 |
| PR | 11-DEC-2000 | 2000US-0256497 |
| PR | 05-JAN-2001 | 2001US-0205908 |

(HITMA -) HITMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-541565/60.

Nucleic acids encoding 3224 human nervous system antigen polypeptides, useful for preventing, diagnosing and/or treating nervous system cancers and metastases -

Claim 1: SEO ID NO 101: 1701pp + Sequence Listing; English.

XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins (ABAI1678-ABAI18001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at <http://wipo.int/pub/publ/abseqs>.

Sequence 305 BP; 72 A; 83 C; 86 G; 63 T; 1 other;

12 19: score 294.6; DB 22; length 305;

```

Best Local Similarity 99.7%; Pred. No. 5.6e-59;
Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

| | | | | |
|---------|------|--------------|----|-----------|
| Matches | 294; | conservative | 27 | 1990-1991 |
|---------|------|--------------|----|-----------|

Query Match 9.8%; Score 237; DB 23; Length 2861;
Best Local Similarity 56.6%; Pred. No. 3.7e-45;
Matches 510; Conservative 0; Mismatches 370; Indels 21; Gaps 3;

1176 TGGCAGACGACCTCATCAACCCCTTTGGAGAGATGATGATTTTGGACCAACTGGA 1235
1 TGGCAGACGACCTCATCAACCCCTTTGGAGAGATGATGATTTTGGACCAACTGGA 60
1236 TTTGCGACAGAAATTTGCAAGTGTCCCTGTTGGCTGTGGATGAGATCCACAGACCTGC 1295
61 TTTGCGACAGAAATTTGCAAGTGTCCCTGTTGGCTGTGGATGAGATCCACAGACCTGC 120
1296 CTCGGATGAGACCCGACATGTAAGTAAGCCGAGCCACAGCCCTTACACAGCTG 1355
121 CTCGGATGAGACCCGACATGTAAGTAAGCCGAGCCACAGCCCTTACACAGCTG 180
1356 CTCGGACCCAGTTCCTGTCGAGCCTCCTTTATGGCTCCACTTCAACATCAGCTGACA 1415
181 CTCGGACCCAGTTCCTGTCGAGCCTCCTTTATGGCTCCACTTCAACATCAGCTGACA 240
1416 AAGAGAGATGAGTTCAGGCCCATCAGAGAGAGAGAGAGATGCTCAGCCTG 1470
241 AAGAGAGATGAGTTCAGGCCCATCAGAGAGAGAGAGAGATGCTCAGCCTG 295

RESULT 12
ABLI0793
ID ABLI0793 standard; cDNA; 2861 BP.
XX ABLI0793;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 26861.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI: 2001-656860/75.
XX P-PSDB: ABB66690.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX Claim 1; SEQ ID NO 26861; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABLI01840-ABLI016175) and the encoded proteins
XX sequences (ABLI01840-ABLI016175).
XX (ABBI7737-ABBI72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 2861 BP; 761 A; 701 C; 723 G; 676 T; 0 other;

Query Match 9.8%; Score 237; DB 23; Length 2861;
Best Local Similarity 56.6%; Pred. No. 3.7e-45;
Matches 510; Conservative 0; Mismatches 370; Indels 21; Gaps 3;

103 CCATGACCATCACTTACACAAACCAAGTGG---CTAATGCCCGCTTAGCCTCTCTCC 159
413 CAAATGACCAATTAAGTACAGAGGTGAAGTGGCCACTTGTGCGCTTTTGTGCA 472
160 GCGTGTGCTGTCGTCGCGGAGAGATCTACAAAGCTGTATATGCGAGTTCTTAATCT 219
473 AATTGCTGCTCAAGATGCGAGAGAGATTTACAACTGTTGGCTAGATCTTCTGCGCT 532
220 TCTGCTGCTCACTACATCATCCTGTTATTTATATAGCTGCGCCCTCAGAGAAACAC 279
533 TCTTGAACATTTACTATGCGATCMACTGATGATGCTTGTGCTCAACCCGACAAA 592
280 AGCTGATGTTGGAAGAACTGATCTGATGACAGCTACATCCAGCTATCCCATTT 339
593 AAGAACTTTGAGGCGCATTTGTCAGTACGTATATGTTACAGAGAACTATACCTCTGT 652
340 CATTGCTGCTGCGCTTCTACGTAAGCTGCTGACCCGCTGCTGTAACCCGTAAGCA 399
653 CATTGCTGCTGCGCTTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 712
400 ACTGCGCGTGGCGCCGACCGGCTCATGAGCCTGCTGCTGCTGCTGCTGCTGCTGCT 459
713 CCATTCCTGCGCGGATCCATCCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 772
460 AGCAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519
773 AGCGAGAGAGCATATGAGGAGGAAATATGATGATGATGATGATGATGATGATGATGAT 832
520 TCTGCGGAGGCTGACACCGCACTCTACAGGCTTCCGAGCGCCGACGACCTGCTG 579
833 TCTGCGGAGGCTTCTGCGGAGGCTGAGGAGGCTTCTGCGGAGGCTTCTGCGGAGGCT 892
580 AAGCAGGCTTATGATGCTCGGACAGCAACAGCAGTGGAGAACTGAGC-----C 630
893 AAGCGGCTGCTGATTAATGATCAATGAAAGACATCATGAGACATGAAAGGCTTTC 952
631 TACACACAGATGTTCTGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
953 CCAGACCTTGAAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012
691 GCGTTGAGAGTTCGATCCGCGGACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750
1013 GAAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1072
751 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 810
1073 AGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1132
811 ATACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
1133 ACACCCAGTGTGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1192
871 AGT-----TTCTGAACCCAGCCAGCCCTACCTGCGCATGAGCTGAGCTGCTG 921
1193 AATGAGCAGATGGAAGAGTGTGGGCAATACATACCTGAAACAGGATGATCTACT 1252
922 TGCCGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 981
1253 TTTCTGATTTACAGCGTGCAGTCTTCTTACATGGGTTGGCTCAAGGTTGGCGGACT 1312
982 C 982
1313 C 1313

RESULT 13
AAV86976
ID AAV86976 standard; cDNA; 424 BP.

Mon Jun 9 10:49:30 2003

Search completed: June 8, 2003, 11:23:51
Job time : 492.416 secs

us-09-622-964-4.rng

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: June 8, 2003, 09:30:48 ; Search time 6124.06 Seconds
(without alignments)
11543.117 Million cell updates/sec

Title: US-09-622-964-4
Sequence: 1 cagggagctccaccagccta.....aaaaaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------------|-----------------------|
| 1 | 2418.4 | 99.6 | 2420 | 9 AF057170 | AF057170 Homo sapi |
| 2 | 1995.4 | 82.1 | 2210 | 9 AF057169 | AF057169 Homo sapi |
| 3 | 1953.2 | 80.4 | 2171 | 9 AF073501 | AF073501 Homo sapi |
| 4 | 648.4 | 26.7 | 1289 | 4 AY064707 | AY064707 Sus scrofa |
| 5 | 642.2 | 26.4 | 142092 | 2 AF139813 | AF139813 Homo sapi |
| 6 | 642.2 | 26.4 | 163915 | 2 AF087451 | AF087451 Homo sapi |
| 7 | 642.2 | 26.4 | 166867 | 9 AC003733 | AC003733 Homo sapi |
| 8 | 642.2 | 26.4 | 196080 | 9 AC004228 | AC004228 Homo sapi |
| 9 | 640.6 | 26.4 | 706 | 9 HSYMD2P10 | AF073499 Homo sapi |
| 10 | 640.6 | 26.4 | 112309 | 9 AC003025 | AC003025 Homo sapi |
| 11 | 640.6 | 26.4 | 133683 | 2 AC084857 | AC084857 Homo sapi |
| 12 | 552.8 | 22.8 | 160169 | 9 AC051664 | AC051664 Homo sapi |
| 13 | 431 | 17.7 | 1908 | 9 AF440756 | AF440756 Homo sapi |
| 14 | 415 | 17.1 | 1956 | 10 BC031186 | BC031186 Mus muscu |
| 15 | 415 | 17.1 | 1957 | 10 BC019528 | BC019528 Mus muscu |
| 16 | 389.2 | 16.0 | 1861 | 9 AK000139 | AK000139 Homo sapi |
| 17 | 358.4 | 14.8 | 492 | 9 HSYMD2P11 | AF073500 Homo sapi |
| 18 | 339.4 | 14.0 | 106648 | 2 AC004588 | AC004588 Homo sapi |
| 19 | 339.4 | 14.0 | 2045 | 9 AF440757 | AF440757 Homo sapi |
| 20 | 338.2 | 13.9 | 1506 | 9 AF440758 | AF440758 Homo sapi |
| 21 | 338.2 | 13.9 | 2500 | 9 AK096459 | AK096459 Homo sapi |
| 22 | 240.8 | 9.9 | 413 | 9 HSYMD2P04 | AF073493 Homo sapi |
| 23 | 240.8 | 9.9 | 1666 | 9 AF052095 | AF052095 Homo sapi |
| 24 | 237 | 9.8 | 2874 | 3 AF218817 | AF218817 Drosophila |
| 25 | 237 | 9.8 | 5610 | 9 AF061546 | AF061546 Drosophila |
| 26 | 232.8 | 9.6 | 295 | 9 HSYMD2P07 | AF073496 Homo sapi |
| 27 | 191.6 | 7.9 | 246 | 10 AF057171 | AF057171 Mus muscu |
| 28 | 189 | 7.8 | 272 | 9 HSYMD2P02 | AF073491 Homo sapi |
| 29 | 189 | 7.8 | 160169 | 2 AC051664 | AC051664 Homo sapi |
| 30 | 188.2 | 7.7 | 32022 | 2 AC019521 | AC019521 Drosophila |
| 31 | 188.2 | 7.7 | 170675 | 3 AC008139 | AC008139 Drosophila |
| 32 | 188.2 | 7.7 | 192055 | 3 AC009183 | AC009183 Drosophila |
| 33 | 188.2 | 7.7 | 221888 | 3 AE003686 | AE003686 Drosophila |
| 34 | 175 | 7.2 | 438 | 11 G05922 | G05922 human SMS W |
| 35 | 175 | 7.2 | 1119 | 9 HUMFERRH | M7166 Human ferritin |
| 36 | 175 | 7.2 | 1198 | 6 AX409527 | AX409527 Sequence |
| 37 | 175 | 7.2 | 1198 | 9 HUMFERRH | L20941 Human ferritin |
| 38 | 161.4 | 6.6 | 239 | 9 HSYMD2P05 | AF073494 Homo sapi |
| 39 | 155 | 6.4 | 125804 | 2 AC019853 | AC019853 Drosophila |
| 40 | 150.4 | 6.2 | 163514 | 3 AC091227 | AC091227 Drosophila |
| 41 | 150.4 | 6.2 | 274351 | 3 AE003531 | AE003531 Drosophila |
| 42 | 150.4 | 6.2 | 274351 | 3 AC018761 | AC018761 Homo sapi |
| 43 | 135.2 | 5.6 | 147750 | 9 AL592166 | AL592166 Human DNA |
| 44 | 130.4 | 5.4 | 90487 | 9 AL611928 | AL611928 Homo sapi |
| 45 | 130.4 | 5.4 | 170141 | 2 | |

ALIGNMENTS

RESULT 1
AF057170 2420 bp mRNA linear PRI 17-OCT-1998
LOCUS Homo sapiens bestrophin (VMD2) mRNA, alternatively spliced product,
complete cds.
VERSION AF057170
ACCESSION AF057170.1 GI:3335160
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2420)
Petrushkin, K., Kolstel, M.J., Bakali, B., Li, W., Xie, G., Marknell, T.,
Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, M.,

Pred. No. 15 is the number of results predicted by chance to have a

Bergen, A.A., McGarty-Dugan, V., Figueroa, D., Austin, C.P.,
 Metzker, M.L., Caskey, C.T. and Wadellius, C.
 Identification of the gene responsible for Best macular dystrophy
 Nat. Genet. 19 (3), 241-247 (1998)
 JOURNAL 9834772
 MEDLINE 9863395
 PUBMED 2 (bases 1 to 2420)
 REFERENCE Petrakian, K.
 AUTHORs Direct Submission
 TITLE Submitted (03-APR-1998) Human Genetics, Merck Research
 JOURNAL Laboratories, West Point, PA 19486, USA
 FEATURES Location/Qualifiers
 source 1..2420
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q13"
 1..2420
 /gene="VMD2"
 105..1412
 /gene="VMD2"
 /note="Best's macular dystrophy gene; BMD; alternatively
 splice isoform; contains exon 7"
 /product="bestrophin"
 /protein_id="A66434.1"
 /db_xref="GI:3353161"
 /translation="MTITTSQVANNRISFRLICWRSIYKLYEFLLECY
 IIRFLYALTEQQLMEKTLTCDYDILIPISYLVGYATLVYTRNNQENLPM
 PDRILSVSGVEKDEQGLRLRLRLANLNLNLSVSTAYVKKRPSNHLVDA
 GPMTPAHKOLEKLSLPHNMFVPMWVFNLSKAWLGGHRLDILQSLINMPLR
 TCGHLAYDMISLPLVYQVTVAVSFLTCLVGRFLNPKATPGLHDLVVEF
 TRDFFYVGMIVKYSRLALGNHGRGHQQLLETMRQCKRYSRVESLMMWT
 VLPATREARAGESLEPGRRLRMQSSSTPLEMMIILPTGLSTGICRCPCMLMR
 CTRTCLMSRTCTGTSHPSPPIOLLPSSVEPPLMAPETSA"
 BASE COUNT 605 a 695 c 612 g 508 t
 ORIGIN
 Query Match 99.6%; Score 2418.4; DB 9; Length 2420;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 CAGGAGTCCACACAGCTAGTCCGACAGCTTCTGTGGATCATCGACCCAGCTGGA 60
 1 CAGGAGTCCACACAGCTAGTCCGACAGCTTCTGTGGATCATCGACCCAGCTGGA 60
 61 CCCACCTGACCAAGCCACCTGCTGAGCCACCTGCTGGCATGACATCACTTACA 120
 61 CCCACCTGACCAAGCCACCTGCTGAGCCACCTGCTGGCATGACATCACTTACA 120
 61 CCCACCTGACCAAGCCACCTGCTGAGCCACCTGCTGGCATGACATCACTTACA 120
 121 CAGGCAAGTGGCTATGCGCTTACGCTCTTCCCGCTGCTGCTGCTGCTGCTG 180
 121 CAGGCAAGTGGCTATGCGCTTACGCTCTTCCCGCTGCTGCTGCTGCTGCTG 180
 121 CAGGCAAGTGGCTATGCGCTTACGCTCTTCCCGCTGCTGCTGCTGCTGCTG 180
 181 GCAGCATCTACAGCTGCTATATGGAGATCTTATCTCTGCTGCTGCTGCTGCTG 240
 181 GCAGCATCTACAGCTGCTATATGGAGATCTTATCTCTGCTGCTGCTGCTGCTG 240
 181 GCAGCATCTACAGCTGCTATATGGAGATCTTATCTCTGCTGCTGCTGCTGCTG 240
 241 TCCGCTTATTTATAGGCTGCGCTCAGGAGAAACAGACGCTGATGTTGAGAACTA 300
 241 TCCGCTTATTTATAGGCTGCGCTCAGGAGAAACAGACGCTGATGTTGAGAACTA 300
 241 TCCGCTTATTTATAGGCTGCGCTCAGGAGAAACAGACGCTGATGTTGAGAACTA 300
 301 CTCTGATTTGAGACAGCTATCATCTATCCCATTTCTCTGCTGCTGCTGCTGCTG 360
 301 CTCTGATTTGAGACAGCTATCATCTATCCCATTTCTCTGCTGCTGCTGCTGCTG 360
 301 CTCTGATTTGAGACAGCTATCATCTATCCCATTTCTCTGCTGCTGCTGCTGCTG 360
 361 TGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 361 TGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 361 TGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 421 TCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 421 TCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 421 TCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480

481 GCAGCTCATTCGCTACGCAACCTGGGACAGTGTCTATCTCTGCGACGCTGACACCG 540
 481 GCAGCTCATTCGCTACGCAACCTGGGACAGTGTCTATCTCTGCGACGCTGACACCG 540
 541 CAGCTACAAAGGCTTCCCGACGCGCCACACCTGTGTGACAGAGCTTTATGATCTCCG 600
 541 CAGCTACAAAGGCTTCCCGACGCGCCACACCTGTGTGACAGAGCTTTATGATCTCCG 600
 541 CAGCTACAAAGGCTTCCCGACGCGCCACACCTGTGTGACAGAGCTTTATGATCTCCG 600
 601 CAGAACACAGAGCTTGGAGAACTAGACCTACCAACAAATGTTCTGGTGCCCTGGG 660
 601 CAGAACACAGAGCTTGGAGAACTAGACCTACCAACAAATGTTCTGGTGCCCTGGG 660
 601 CAGAACACAGAGCTTGGAGAACTAGACCTACCAACAAATGTTCTGGTGCCCTGGG 660
 661 TGTGCTTCCCAACCTGTCAATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 661 TGTGCTTCCCAACCTGTCAATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 661 TGTGCTTCCCAACCTGTCAATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 721 TGTCTCAGAGCTCTCTAAGAGATGAAACCTTGTGCTGCTGCTGCTGCTGCTGCTG 780
 721 TGTCTCAGAGCTCTCTAAGAGATGAAACCTTGTGCTGCTGCTGCTGCTGCTGCTG 780
 721 TGTCTCAGAGCTCTCTAAGAGATGAAACCTTGTGCTGCTGCTGCTGCTGCTGCTG 780
 781 CCTACGAGCTGATTTGATATCCCACTGGTATACAGAGTGTGCTGCTGCTGCTGCTG 840
 781 CCTACGAGCTGATTTGATATCCCACTGGTATACAGAGTGTGCTGCTGCTGCTGCTG 840
 781 CCTACGAGCTGATTTGATATCCCACTGGTATACAGAGTGTGCTGCTGCTGCTGCTG 840
 841 GCTTCTCTGACTTGTCTAGTTGGGCGGAGTTTCTGACACCCAGCCAGGCTTACCTG 900
 841 GCTTCTCTGACTTGTCTAGTTGGGCGGAGTTTCTGACACCCAGGCTTACCTGCTG 900
 841 GCTTCTCTGACTTGTCTAGTTGGGCGGAGTTTCTGACACCCAGGCTTACCTGCTG 900
 901 GCGATGAGCTGAGCTCTGTTGCGCCGCTTACAGTTCTGCTGCTGCTGCTGCTGCTG 960
 901 GCGATGAGCTGAGCTCTGTTGCGCCGCTTACAGTTCTGCTGCTGCTGCTGCTGCTG 960
 901 GCGATGAGCTGAGCTCTGTTGCGCCGCTTACAGTTCTGCTGCTGCTGCTGCTGCTG 960
 961 GCTGGGTGAAGTGGGCGCTCTCAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 961 GCTGGGTGAAGTGGGCGCTCTCAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 961 GCTGGGTGAAGTGGGCGCTCTCAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 1021 ATGGCCAGAGCTGCTTGAAGAGAGATGATGTCAGAGAAAGAGGCTGACGCTGAG 1080
 1021 ATGGCCAGAGCTGCTTGAAGAGAGATGATGTCAGAGAAAGAGGCTGACGCTGAG 1080
 1021 ATGGCCAGAGCTGCTTGAAGAGAGATGATGTCAGAGAAAGAGGCTGACGCTGAG 1080
 1081 AAGGAG 1140
 1081 AAGGAG 1140
 1081 AAGGAG 1140
 1141 GAGAAATCGCTTGAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 1141 GAGAAATCGCTTGAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 1141 GAGAAATCGCTTGAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 1201 TGGAGAGAGATGATGATTTTGAAGCAACAGTGTGTCAGAGAGAGATTTGAGAGT 1260
 1201 TGGAGAGAGATGATGATTTTGAAGCAACAGTGTGTCAGAGAGAGATTTGAGAGT 1260
 1201 TGGAGAGAGATGATGATTTTGAAGCAACAGTGTGTCAGAGAGAGATTTGAGAGT 1260
 1261 TGGAGAGAGATGATGATTTTGAAGCAACAGTGTGTCAGAGAGAGATTTGAGAGT 1320
 1261 TGGAGAGAGATGATGATTTTGAAGCAACAGTGTGTCAGAGAGAGATTTGAGAGT 1320
 1261 TGGAGAGAGATGATGATTTTGAAGCAACAGTGTGTCAGAGAGAGATTTGAGAGT 1320
 1321 GAATTAAGCCGAG 1380
 1321 GAATTAAGCCGAG 1380
 1321 GAATTAAGCCGAG 1380
 1381 CTTTATGAGCTTCAACCTTCAACATCAGAGCTGAAACAAAGAGAGATGAGTTCAG 1440
 1381 CTTTATGAGCTTCAACCTTCAACATCAGAGCTGAAACAAAGAGAGATGAGTTCAG 1440
 1381 CTTTATGAGCTTCAACCTTCAACATCAGAGCTGAAACAAAGAGAGATGAGTTCAG 1440
 1441 TCAG 1500
 1441 TCAG 1500
 1441 TCAG 1500
 1501 CCATGATACCATCTCTCCAGAGGCAAACTCAAGAGCAAACTACTGTTGGCCAAAGAG 1560
 1501 CCATGATACCATCTCTCCAGAGGCAAACTCAAGAGCAAACTACTGTTGGCCAAAGAG 1560
 1501 CCATGATACCATCTCTCCAGAGGCAAACTCAAGAGCAAACTACTGTTGGCCAAAGAG 1560

QY 1561 ATCCCTCTCCAGAGGCGCTGCCAAAACACAGGACCCAAACAGAACCTTAGGG 1620
 DB 1561 ATCCCTCTCCAGAGGCGCTGCCAAAACACAGGACCCAAACAGAACCTTAGGG 1620
 QY 1621 CCAGGAAGACAAAGGCGCTGGAGCTTAAGGCTGTGAGAGCTTCAAGTCTGCCACT 1680
 DB 1621 CCAGGAAGACAAAGGCGCTGGAGCTTAAGGCTGTGAGAGCTTCAAGTCTGCCACT 1680
 QY 1681 GTATCAGAGGCGAGCTACTAGTGCCTCCACAGAGCGCCCTCAGCCCTCCCATGTT 1740
 DB 1681 GTATCAGAGGCGAGCTACTAGTGCCTCCACAGAGCGCCCTCAGCCCTCCCATGTT 1740
 QY 1741 CTTCCTCCCTAGAACATCAGGCGCGCTCAAGCTTCAAGCTTCAAGGACATAGACCA 1800
 DB 1741 CTTCCTCCCTAGAACATCAGGCGCGCTCAAGCTTCAAGCTTCAAGGACATAGACCA 1800
 QY 1801 AGACAAAAGCTTAAAGACTGTAGTCTGGGGCCCAAGAAAGTTTGAATGCTCTAGA 1860
 DB 1801 AGACAAAAGCTTAAAGACTGTAGTCTGGGGCCCAAGAAAGTTTGAATGCTCTAGA 1860
 QY 1861 GAGCGATGGGCGCTTGTATGAGACACCCAGAGATCTCAAGTGTAGAGAGAAAAGTGTGA 1920
 DB 1861 GAGCGATGGGCGCTTGTATGAGACACCCAGAGATCTCAAGTGTAGAGAGAAAAGTGTGA 1920
 QY 1921 GTTTAACTGACGAGTATGCCAGAGATCCCGAAATCAGCTCAAGAACCTTTGAGACA 1980
 DB 1921 GTTTAACTGACGAGTATGCCAGAGATCCCGAAATCAGCTCAAGAACCTTTGAGACA 1980
 QY 1981 ATCACCACCAACATACACACTACATCACTCAAGATCAGATGATGCTTATGGGCTTGA 2040
 DB 1981 ATCACCACCAACATACACACTACATCACTCAAGATCAGATGATGCTTATGGGCTTGA 2040
 QY 2041 AAACAGGAGTGAAGCATCTTCTAACCTCTCCAAATGGGATGCTTCCAGGACAGGT 2100
 DB 2041 AAACAGGAGTGAAGCATCTTCTAACCTCTCCAAATGGGATGCTTCCAGGACAGGT 2100
 QY 2101 CCTCAGCTGTGTACAGCAGGACGACATGATCCAGTACAGACCATACAGCTGTCCACA 2160
 DB 2101 CCTCAGCTGTGTACAGCAGGACGACATGATCCAGTACAGACCATACAGCTGTCCACA 2160
 QY 2161 CTGAGAGACGTGTCTACACAGCCTGGAATCAATGGTATAGTATATAAATATCC 2220
 DB 2161 CTGAGAGACGTGTCTACACAGCCTGGAATCAATGGTATAGTATATAAATATCC 2220
 QY 2221 AGACACTCTCAGGCTTAAATGCTTATATCAATGTAAGTGAAGTGAACCA 2280
 DB 2221 AGACACTCTCAGGCTTAAATGCTTATATCAATGTAAGTGAAGTGAACCA 2280
 QY 2281 TTGGAACAATTAATCAGACTGTGATTCAGAGTGGGAACTTATGTTCTGGAAT 2340
 DB 2281 TTGGAACAATTAATCAGACTGTGATTCAGAGTGGGAACTTATGTTCTGGAAT 2340
 QY 2341 CCAAGACAGCAGACCTTACTATAGCCCAAACTAATGATTAAATACAAATACT 2400
 DB 2341 CCAAGACAGCAGACCTTACTATAGCCCAAACTAATGATTAAATACAAATACT 2400
 QY 2401 CGTTAAAAAATAAAAA 2420
 DB 2401 CGTTAAAAAATAAAAA 2420

RESULT 2
 AF057169 2210 bp mRNA linear PRI 17-OCT-1998
 LOCUS Homo sapiens beetrophin (VMD2) mRNA, alternatively spliced product,
 DEFINITION complete cds.
 AF057169
 VERSION AF057169.1 GI:3335158
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2210)
 AUTHORS Petrakikh, K., Kozel, M. J., Bakall, B., Li, W., Xie, G., Marknell, T.,
 Sandgren, O., Forsman, K., Holmgren, G., Andersson, S., Vujic, M.,
 Bergen, A. A., McGarity-Dugan, V., Figueroa, D., Austin, C. P.,
 Metzger, M. L., Caskey, C. T. and Madeline, C.
 TITLE Identification of the gene responsible for Best macular dystrophy
 JOURNAL Nat. Genet. 19 (3), 241-247 (1998)
 MEDLINE 98324772
 PUBMED 9662395
 REFERENCE 2 (bases 1 to 2210)
 AUTHORS Petrakikh, K.
 TITLE Direct Submission
 JOURNAL Submitted (03-APR-1998) Human Genetics, Merck Research
 Laboratories, West Point, PA 19486, USA
 FEATURES
 source
 1. 2210
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q13"
 1. 2210
 /gene="VMD2"
 105..1862
 /note="Best's macular dystrophy gene; BMD; alternatively
 spliced isoform"
 /product="bestrophin"
 /protein_id="AAC64343.1"
 /db_xref="GI:3335159"
 /translation="MTTPTYSOVNANRLAGSRLLLCMRISYKLYGEFLILCY
 IIRITRYALTEBOQLMPKILTXCDXYIOLIPSYLGYVTLVYTRMNOENHVOA
 PDRLMSLVSGVEGKDEQGLRLRLTIRLAKGNVLILBSYSTRVRRPSNOYLVA
 GFMTPAERKOLEKLSLPHNMFWVWVFWANLSKAWLGGTRIDILQSILNENRTR
 TQCHFLAYDWISLPLVTOVVVAVSVFELTCLVGNQFLNPAKATPGEHDLVYVF
 TLOFFEVYGLKVAEOLINFGEDDDFFETNIVDNLLVSLAVDEHODLPRME
 DMVWNRKPPPTVAAASQFRAFGSGSTENISLKEEMEFOPNDEDEAHAGILGR
 FICGSHDHHRPRANSRTRKILMPRESLHGLRKNHKAQNVGQEDNKAARLKA
 DAFSAPIYORPGYASAPQPLSPTPMPEFLPSPAPSLHSVTGIDTRKDSLKTVSSG
 AKKSEFLSBDGALMEHPYVQYRKRIVERNLDMELIPNHLKEPLDEGSPNTHRT
 LKDHMDPYMALENDEAHS"
 BASE COUNT 556 a 647 c 531 g 476 t
 ORIGIN
 Query Match 82.1%; Score 1995.4; DB 9; Length 2210;
 Best Local Similarity 91.5%; Pred. No. 0; Mismatches 1; Indels 203; Gaps 1;
 Matches 2209; Conservative 0;

QY 1 CAGGAGTCCACACAGCTAGTCCAGACCTTCTGTGGATCATGAGCCACCTGGA 60
 DB 1 CAGGAGTCCACACAGCTAGTCCAGACCTTCTGTGGATCATGAGCCACCTGGA 60
 QY 61 CCCCACTGACCCAGCCACCTGCTGCAGCCACATGCTGCGCATGACATCATCA 120
 DB 61 CCCCACTGACCCAGCCACCTGCTGCAGCCACATGCTGCGCATGACATCATCA 120
 QY 121 CAAGCAATGGCTATGCGCGCTTACGCTTCCCGCTGCTGCTGCTGCGCG 180
 DB 121 CAAGCAATGGCTATGCGCGCTTACGCTTCCCGCTGCTGCTGCTGCGCG 180
 QY 181 GCAGCATCTACAGCTGTATATGAGGAGTCTTAACTCTCTGCTCTGCTACTAC 240
 DB 181 GCAGCATCTACAGCTGTATATGAGGAGTCTTAACTCTCTGCTCTGCTACTAC 240
 QY 241 TCCGCTTATTTATAGGCGGCGCCACAGGAAGAACACAGCTGATGTTGAGAACTA 300
 DB 241 TCCGCTTATTTATAGGCGGCGCCACAGGAAGAACACAGCTGATGTTGAGAACTA 300
 QY 301 CTCTGATATGCGACACTACATCCAGCTATCCCATTTCTTCTGCTGCTGCTTACG 360
 DB 301 CTCTGATATGCGACACTACATCCAGCTATCCCATTTCTTCTGCTGCTGCTTACG 360
 QY 361 TGAAGCTGTGCTGACCCGCTGTGGAACACAGTACAGAACTGCCGTGGCCGACCGCC 420

Db 361 TGACGCTGGGTCTGACCCCGCTGGTAGACCAAGTACAGAACTGCCGTGCCCCACCC 420
 Oy 421 TCATGAGCCGTGTGTGCGGGCTTGTGTGAAGGCAAGGACAGCAAGCGCGCTGTGCGC 480
 Db 421 TCATGAGCCGTGTGTGCGGGCTTGTGTGAAGGCAAGGACAGCAAGCGCGCTGTGCGC 480
 Oy 481 GCACGCTCATCCGCTACGCCACCTGGGCAACGTGCTCATCTGGCGACGTCACACCG 540
 Db 481 GCACGCTCATCCGCTACGCCACCTGGGCAACGTGCTCATCTGGCGACGTCACACCG 540
 Oy 541 CAGCTCAACAGCGCTTCCCGAGGCCAGCACTGGTGCAGCAAGCGCTTATGACTCGG 600
 Db 541 CAGCTCAACAGCGCTTCCCGAGGCCAGCACTGGTGCAGCAAGCGCTTATGACTCGG 600
 Oy 601 CAGAACCAAGCAATTTGGAGAACTGACCTACACACAACATGTTCTGGGTGCCCTGG 660
 Db 601 CAGAACCAAGCAATTTGGAGAACTGACCTACACACAACATGTTCTGGGTGCCCTGG 660
 Oy 661 TGTGTTTGGCCAACTGTCAATGAAGCGTGGCTGGAGGTGCATACCGGAGCCATATC 720
 Db 661 TGTGTTTGGCCAACTGTCAATGAAGCGTGGCTGGAGGTGCATACCGGAGCCATATC 720
 Oy 721 TGTCTCAGAGCCTGCTGMAACGAGTGAACACCTTGCTACTCAGTGTGACACTGTATG 780
 Db 721 TGTCTCAGAGCCTGCTGMAACGAGTGAACACCTTGCTACTCAGTGTGACACTGTATG 780
 Oy 781 CCTACGACTGGATTAGTATCCACTGGGTATACAGAGGTGGAGACTGTGGCGGTATAC 840
 Db 781 CCTACGACTGGATTAGTATCCACTGGGTATACAGAGGTGGAGACTGTGGCGGTATAC 840
 Oy 841 GCTTCTCTGCTGACTGTGTAGTTGGGGGGCAGTTTCTGAACCCAGCCAGGCTTACC 900
 Db 841 GCTTCTCTGCTGACTGTGTAGTTGGGGGGCAGTTTCTGAACCCAGCCAGGCTTACC 900
 Oy 901 GCCATGAGCTGGACCTCGTTGGCCCGGCTCAAGTTCCTGACGTTCTTCTCATGTTG 960
 Db 901 GCCATGAGCTGGACCTCGTTGGCCCGGCTCAAGTTCCTGACGTTCTTCTCATGTTG 960
 Oy 961 GCTGCTGMAAGGTGGGCGCTCTCAAGGCGCTGCTGGGCTGGAGGCTATGCCAGAGG 1020
 Db 961 GCTGCTGMAAGGTGGGCGCTCTCAAGGCGCTGCTGGGCTGGAGGCTATGCCAGAGG 1020
 Oy 1021 ATGGCCAGCAGCTGCTTGAGACGAGAGTGCAGTGTCAAGAAAGGAAGTCTCACGG 1080
 Db 971 ----- 970
 Oy 1081 AAGCAGCCAGCGGTGGTGGCGCACACCTGTAACTCCAGTACTCGGAGGCTGAGCG 1140
 Db 971 ----- 970
 Oy 1141 GAGAAATCGCTTGAACCCGGGAGCGGAGGTGTGTGTGCAGACGCTCATCAACCC 1200
 Db 971 -----GGTGGCAGAGCGCTCATCAACCCCTT 997
 Oy 1201 TGGAGAGGATGATGATTTTGAAGCCAACTGATGTGCAGACGAATTTGCAAGTGC 1260
 Db 998 TGGAGAGGATGATGATTTTGAAGCCAACTGATGTGCAGACGAATTTGCAAGTGC 1057
 Oy 1261 CCGTGTGGCTGTGGATGAGATGCACAGGACCTGCTCGGATGGAAGCCGAGATGTACTG 1320
 Db 1058 CCGTGTGGCTGTGGATGAGATGCACAGGACCTGCTCGGATGGAAGCCGAGATGTACTG 1117
 Oy 1321 GAATTAAGCCGAGACCAAGCCGCCCTTAACAAGTGTTCGCGCCAGTTCCGTCGAGCTC 1380
 Db 1118 GAATTAAGCCGAGACCAAGCCGCCCTTAACAAGTGTTCGCGCCAGTTCCGTCGAGCTC 1177
 Oy 1381 CTTTATGGCTCCACCTTCAACATCAGCCTTGAAACAAAGAGAGATGGACTTCAGCCCA 1440
 Db 1178 CTTTATGGCTCCACCTTCAACATCAGCCTTGAAACAAAGAGAGATGGACTTCAGCCCA 1237
 Oy 1441 TCAGAGAGCAGAGAGATGCTACGCTGGCATCATTTGGCCGTTCTTAGGCTCGAATC 1500

| Db | 1238 | TCAGGAGACGAGAGAGATGCTCAGCCTGGCATCATTTGGCCCTTCCTAGAGCTCGACTC | 1297 |
|----|------|---|------|
| QY | 1501 | CAATGATCACCATCTCTCCAGGGCAAACTCAAGACCAACTACTGTGGCCCAAGAGGA | 1560 |
| Db | 1298 | CCAATGATCACCATCTCTCCAGGGCAAACTCAAGACCAACTACTGTGGCCCAAGAGGA | 1357 |
| QY | 1561 | ATCCCTTCTCCAGAGGGCTGCCCAAAAACCAAGGAGGCAAAAGAAAGCTTAGAGG | 1620 |
| Db | 1358 | ATCCCTTCTCCAGAGGGCTGCCCAAAAACCAAGGAGGCAAAAGAAAGCTTAGAGG | 1417 |
| QY | 1621 | CCAGGAAGACAAGGCGCTGGAAGCTTAAGGCTGTGAGAGGCTTCAGTCTGGCCACT | 1680 |
| Db | 1418 | CCAGGAAGACAAGGCGCTGGAAGCTTAAGGCTGTGAGAGGCTTCAGTCTGGCCACT | 1477 |
| QY | 1681 | GTATCAGAGGCCAGGCTACTACAGTGGCCCAAGAGCGCCCTCAGCCCACTGCCATGTT | 1740 |
| Db | 1478 | GTATCAGAGGCCAGGCTACTACAGTGGCCCAAGAGCGCCCTCAGCCCACTGCCATGTT | 1537 |
| QY | 1741 | CTTCCCTCAGACCATCAGCGCGGTAAAGCTTCAAGTGTACAGGGCATAGACACAA | 1800 |
| Db | 1538 | CTTCCCTCAGACCATCAGCGCGGTAAAGCTTCAAGTGTACAGGGCATAGACACAA | 1597 |
| QY | 1801 | AGACAAAAGCTTAAAGACTGTGAGTCTGGGGCCAGAAAAAGTTTGAATTGCTCTCAG | 1860 |
| Db | 1598 | AGACAAAAGCTTAAAGACTGTGAGTCTGGGGCCAGAAAAAGTTTGAATTGCTCTCAG | 1657 |
| QY | 1861 | GAGCGATGGGCGCTTGATGAGACACCCAGAAATATCTCAAGTGTAGAGGAAACGTGG | 1920 |
| Db | 1658 | GAGCGATGGGCGCTTGATGAGACACCCAGAAATATCTCAAGTGTAGAGGAAACGTGG | 1717 |
| QY | 1921 | GTTTAACCTGAGGGATATGCGAGAGATCCCGAAATACACCTCAAGAACCTTTGGAA | 1980 |
| Db | 1718 | GTTTAACCTGAGGGATATGCGAGAGATCCCGAAATACACCTCAAGAACCTTTGGAA | 1777 |
| QY | 1981 | ATCACCAACAACATACACTACACTCAAAAGATCACATGATCCTTATTTGGGCTTGG | 2040 |
| Db | 1778 | ATCACCAACAACATACACTACACTCAAAAGATCACATGATCCTTATTTGGGCTTGG | 1837 |
| QY | 2041 | AAACAGGATGAAGACATCTCTTAACCTGCTTCTTAATGGGATGCTTGGCAGCCAGGT | 2100 |
| Db | 1838 | AAACAGGATGAAGACATCTCTTAACCTGCTTCTTAATGGGATGCTTGGCAGCCAGGT | 1897 |
| QY | 2101 | CCGACCTGTGTACACAGAGACACTATCCAGTCAAGCCATACAGCTGTCCACA | 2160 |
| Db | 1898 | CCGACCTGTGTACACAGAGACACTATCCAGTCAAGCCATACAGCTGTCCACA | 1957 |
| QY | 2161 | CTGNAAGACGTGTCTTACAAACAGCTGATCAATGTTAGCTTAATGATAAAAATCCC | 2220 |
| Db | 1958 | CTGNAAGACGTGTCTTACAAACAGCTGATCAATGTTAGCTTAATGATAAAAATCCC | 2017 |
| QY | 2221 | AGACTACTTCAGCCTTAAATGCGTTTATTCATTAATAAAGCTGGAAGCTAGTGAACA | 2280 |
| Db | 2018 | AGACTACTTCAGCCTTAAATGCGTTTATTCATTAATAAAGCTGGAAGCTAGTGAACA | 2077 |
| QY | 2281 | TTGGAACAATTTAACTCAGACTCTGTGATTCAGAGTCTGGAAACCTTACTTATCTGAT | 2340 |
| Db | 2078 | TTGGAACAATTTAACTCAGACTCTGTGATTCAGAGTCTGGAAACCTTACTTATCTGAT | 2137 |
| QY | 2341 | CCAAGACAGCCACACTTATGATTAAGTCCCAAACTAATGAGTTTAATAAACAATAACT | 2400 |
| Db | 2138 | CCAAGACAGCCACACTTATGATTAAGTCCCAAACTAATGAGTTTAATAAACAATAACT | 2197 |
| QY | 2401 | CGTTAAAAAAA 2413 | |
| Db | 2198 | CGTTAAAAAAA 2210 | |

| | |
|------------|---|
| RESULT 3 | |
| AF073501 | 2171 bp mRNA linear PRI 02-SEP-1996 |
| LOCUS | |
| DEFINITION | Homo sapiens vitelliform macular dystrophy protein (VMD2) mRNA, |
| | complete cds. |
| ACCESSION | AF073501 |

VERSION AF073501.1 GI:3511241
 KEYWORDS Homo sapiens.
 SOURCE Homo sapiens.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1. (bases 1 to 2171)
 AUTHORS Stohr, H., Marguardt, A., Rivera, A., Cooper, P.R., Nowak, N.J., Shows, T.B., Gerhard, D.S. and Weber, B.H.
 TITLE A gene map of the Best's vitelliform macular dystrophy region in chromosome 11q12-q13.1
 JOURNAL Genome Res. 8 (1), 48-56 (1998)
 MEDLINE 98112782
 PUBMED 9445487
 REFERENCE 2. (bases 1 to 2171)
 AUTHORS Marguardt, A., Stohr, H., Passmore, L., Kraemer, F., Rivera, A. and Weber, B.H.
 TITLE Direct Submission
 JOURNAL Submitted (22-JUN-1998) Human Genetics, University, Blozentrum, Am Hubland, Wurzburg 97074, Germany
 FEATURES
 source
 1. 2171
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q12-q13.1"
 /tissue_type="retinal pigment epithelium"
 1. 2171
 /gene="VMD2"
 89..1846
 /gene="VMD2"
 /note="Best disease"
 /codon_start=1
 /product="vitelliform macular dystrophy protein"
 /protein_id="AAC33766.1"
 /db_xref="GI:3511242"
 /translation="MTITVSVANARLGSFRLLCWNGSTYKLLYGEPLIFLICY IREFIYRLAEQQLMEKLYCDSTYOLIPISFVLEFVTLVTVRWKQENLVP PRLMSLVGFEQKDEQRLRLRLRYANGLNVLRSVSTAVYKRPSQHLVQA GKTPRAEHLKELISLPHMFWPMWFMNLSKMAVGRIDPILLOSLNEMNTLR TQCGHLKATDKLISPLVYQVTVAVSFPLICVGRPLNPAKAVPGEHLVYVE FLQGFYVGMKLVAVQLINPREDDEDETRNIYNDIYSLVADENHODLPHEP DMYWNKPEQPPYTAASQFRASFEWGSFNFSLNKEENEPQNDDEDAAGIIGR FLGLQSHDHPRPANRSRTKLMPKRESLHEGLPKNHKAKOVNROEDKMKMLKAV DAKSAPLYORPQYSAQPLSPTRMPEPLPSAPSKLSVYGLDTPKXSLKTVSSG AKKSPFLSDGALMEHPEVSOVRRKTEFNLTPMEIPEHLKEPLBROSPTNHTT LKDHMPYALLENRDEAHS"
 BASE COUNT 540 a 638 c 523 g 470 t
 ORIGIN
 Query Match 80.4%; Score 1953.2; DB 9; Length 2171;
 Best Local Similarity 91.3%; Pred. No. 0;
 Matches 2168; Conservative 0; Mismatches 3; Indels 203; Gaps 1;

QY 317 CTACATCCAGCTCATCCCATTTCTTCTGCTGGCTTCACTGACGCTGCTGAC 376
 DB 301 CTACATCCAGCTCATCCCATTTCTTCTGCTGGCTTCACTGACGCTGCTGAC 360
 QY 377 CCGCTGGTGGAAACGATACAGAACCTGCGCGCCGACCGCCATAGAGCTGTGTC 436
 DB 361 CCGCTGGTGGAAACGATACAGAACCTGCGCGCCGACCGCCATAGAGCTGTGTC 420
 QY 437 GGGCTTCGTCGAAGCAAGAGACAGAGCGCGCTGTGCGGCGACGCTCATCCGCTA 496
 DB 421 GGGCTTCGTCGAAGCAAGAGACAGAGCGCGCTGTGCGGCGACGCTCATCCGCTA 480
 QY 497 CGCCAACTGGGCAAGCTCTATCTCGCGACGCTCAGACCGCAAGTCTAAGCGCT 556
 DB 481 CGCCAACTGGGCAAGCTCTATCTCGCGACGCTCAGACCGCAAGTCTAAGCGCT 540
 QY 557 CCCGACGCGCCAGACCTGCTGCAACGAGCGTTATGATCCGCGAGAACACAGAGTT 616
 DB 541 CCCGACGCGCCAGACCTGCTGCAACGAGCGTTATGATCCGCGAGAACACAGAGTT 600
 QY 617 GGAGAACTGAGCTTACACACACACATGTTCTGGTGCCCTGGGTGTGTTGCCACT 676
 DB 601 GGAGAACTGAGCTTACACACACACATGTTCTGGTGCCCTGGGTGTGTTGCCACT 660
 QY 677 GTCAATGAGCGCTGCTTGGAGTGGAAATCCGGGACCTATCTCTCAGAGCTGCT 736
 DB 661 GTCAATGAGCGCTGCTTGGAGTGGAAATCCGGGACCTATCTCTCAGAGCTGCT 720
 QY 737 GAACGATGAAACACTTGGCTACTAGTGTGACACCTGATGCTTACGACTGATTAAG 796
 DB 721 GAACGATGAAACACTTGGCTACTAGTGTGACACCTGATGCTTACGACTGATTAAG 780
 QY 797 TATCCACTGCTATATACAGAGTGTGATGTGGCGGTGTACAGCTTCTCTGACTTG 856
 DB 781 TATCCACTGCTATATACAGAGTGTGATGTGGCGGTGTACAGCTTCTCTGACTTG 840
 QY 857 TCTAGTGGGGGAGGATTTGAAACCGACGCAAGGCTTACCTGGCATAGCTGACCT 916
 DB 841 TCTAGTGGGGGAGGATTTGAAACCGACGCAAGGCTTACCTGGCATAGCTGACCT 900
 QY 917 CGTTGTGCGCTCTTACAGTTCCTGACAGTTCCTTATGTTGGCTGGAAGTGCGG 976
 DB 901 CGTTGTGCGCTCTTACAGTTCCTGACAGTTCCTTATGTTGGCTGGAAGTGCGG 954
 QY 977 CCTCTCAGAGGCGCTGCTGAGGCTGAGGACATGGCAGAGGCTATGCGCAGAGCTGCT 1036
 DB 955 ----- 954
 QY 1037 TGAGACGAGATGATGCTGACAGAAAGAGGCTCACGGGTAGAAAGACCGCAGCGTG 1096
 DB 955 ----- 954
 QY 1097 GTGGCGACACTGTAAATCCAGCTATCTCGGAGGCTGAGGACAGAAATGCTGAACC 1156
 DB 955 ----- 954
 QY 1157 CGGAGAGCGAGTGTGGTGACAGAGCTCATCAACCCCTTTGAGAGATGATGAT 1216
 DB 955 ----- 954
 QY 1217 GATTTGAGACCACTGATTTGACAGAGATTTGAGAGTTCCTGCTGCTGAT 1276
 DB 998 GATTTGAGACCACTGATTTGACAGAGATTTGAGAGTTCCTGCTGCTGAT 1057
 QY 1277 GAGATGACAGAGCTGCTGAGTGTGAGAGCGGACATGTACTGGAATTAAGCCGAGCA 1336
 DB 1058 GAGATGACAGAGCTGCTGAGTGTGAGAGCGGACATGTACTGGAATTAAGCCGAGCA 1117
 QY 1337 CAGCCCGCTACAGAGCTGCTGCGCCAGTTCGCTCAGAGCTCTTATAGGGCTCAGC 1396
 DB 1118 CAGCCCGCTACAGAGCTGCTGCGCCAGTTCGCTCAGAGCTCTTATAGGGCTCAGC 1177

1397 TTCAACATCAGCCGTAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGAGAGAG 1456
 1178 TTCAACATCAGCCGTAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGAGAGAG 1237
 1457 GATGCTCAGCGTGGATGATGAGCGGCTTCTAGGCTGAGTCCATGATCACCATT 1516
 1238 GATGCTCAGCGTGGATGATGAGCGGCTTCTAGGCTGAGTCCATGATCACCATT 1297
 1517 CCCAGGCAAACTCAAGAGCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCCAG 1576
 1298 CCCAGGCAAACTCAAGAGCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCCAG 1357
 1577 GGCCTGCCCAAAACCAAGGACCAACAGACGTTAGGGGCGAGAGCAAAAG 1636
 1358 GGCCTGCCCAAAACCAAGGACCAACAGACGTTAGGGGCGAGAGCAAAAG 1417
 1637 GGCCTGCCCAAACTCAAGGACCAACAGACGTTAGGGGCGAGAGCAAAAG 1696
 1418 GGCCTGCCCAAACTCAAGGACCAACAGACGTTAGGGGCGAGAGCAAAAG 1477
 1697 TACTACAGTCCGACAGACGCGCTCAGCCGACCTCCATGTTCTTCCCTAGAAC 1756
 1478 TACTACAGTCCGACAGACGCGCTCAGCCGACCTCCATGTTCTTCCCTAGAAC 1537
 1757 TCAAGCCGCTCAAAAGCTTCAAGTGTCAAGGACATGACCAAAAGAAAGCTTAA 1816
 1538 TCAAGCCGCTCAAAAGCTTCAAGTGTCAAGGACATGACCAAAAGAAAGCTTAA 1597
 1817 ACTGTAGTCTGGGCGCAAGAAAGTTTGAATGCTCTCAGAGACGATGGGCGCTG 1876
 1598 ACTGTAGTCTGGGCGCAAGAAAGTTTGAATGCTCTCAGAGACGATGGGCGCTG 1657
 1877 ATGAGACCCCAAGATATCTCAAGTGAAGAGAGAAACGTGGAGTTTAACTGACGAT 1936
 1658 ATGAGACCCCAAGATATCTCAAGTGAAGAGAGAAACGTGGAGTTTAACTGACGAT 1717
 1937 ATGAGACCCCAAGATATCTCAAGTGAAGAGAGAAACGTGGAGTTTAACTGACGAT 1996
 1718 ATGAGACCCCAAGATATCTCAAGTGAAGAGAGAAACGTGGAGTTTAACTGACGAT 1777
 1997 CACACTACACTCAAAAGATCAGATGATCCTTATTTGGGCTTGGAAACAGAGATGAGCA 2056
 1778 CACACTACACTCAAAAGATCAGATGATCCTTATTTGGGCTTGGAAACAGAGATGAGCA 1837
 2057 CATCTCAACCTGCTTCTTAATGGGAGTCTCCGACGACGATCTCAGCTGTGTATC 2116
 1838 CATCTCAACCTGCTTCTTAATGGGAGTCTCCGACGACGATCTCAGCTGTGTATC 1897
 2117 ACCAGCAGACAGTATCAGATCAGACGATCAGCTCTCAGACTGAAAGAGTGTCT 2176
 1898 ACCAGCAGACAGTATCAGATCAGACGATCAGCTCTCAGACTGAAAGAGTGTCT 1957
 2177 ACCAGCAGTAAATGATGTTAGCTTAATGATTAATGATTAATGATTAATGATTAATGAT 2236
 1958 ACCAGCAGTAAATGATGTTAGCTTAATGATTAATGATTAATGATTAATGATTAATGAT 2017
 2237 TAATGCTTTTATTAATAAACTGTGAAGCTAGACTGAAACATTTAACT 2296
 2018 TAATGCTTTTATTAATAAACTGTGAAGCTAGACTGAAACATTTAACT 2077
 2297 CAGACTGTGATTCAGAGTCCGGGAACTTATGTTATCTGAAATCCAGAGCAGCACC 2356
 2078 CAGACTGTGATTCAGAGTCCGGGAACTTATGTTATCTGAAATCCAGAGCAGCACC 2137
 2357 TTAGATATGCTGCCCAAACTAATGATTAATAAA 2390
 2138 TTAGATATGCTGCCCAAACTAATGATTAATAAA 2171

RESULT 4
 LOCUS AY064707 1289 bp mRNA linear MM 19-AUG-2002
 DEFINITION Sus scrofa domestica mRNA, partial cds.

ACCESSION AY064707
 VERSION AY064707.1 GI:18476495
 KEYWORDS
 SOURCE
 ORGANISM
 Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 Marmorstein, L.Y., McLaughlin, P.J., Stanton, J.B., Yan, L., Crabbs, J.W.
 and Marmorstein, A.D.
 Bestrophin interacts physically and functionally with protein
 phosphatase 2A
 J. Biol. Chem. 277 (34), 30591-30597 (2002)
 12058047
 2 (bases 1 to 1289)
 Marmorstein, L.Y., McLaughlin, P.J., Stanton, J.B., Yan, L., Crabbs, J.W.
 and Marmorstein, A.D.
 Direct Submission
 Submitted (30-NOV-2001) Cole Eye Institute, 131, Cleveland Clinic
 Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA
 FEATURES
 source
 1. 1289
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 <1..1289
 /codon_start=3
 /product="bestrophin"
 /protein_id="AA040882.1"
 /db_xref="GI:18476495"
 /translation="PQHLYKAGFMTSPSEKHLEKSLPHNSFMRPWYFANLSIKAWI
 GRRIDPVLQSLDEMTLTPROCHLAVYMTISVPLTVTVVAVYSFLACVGR
 OFNLAKVYVGHMDLVLPFLFFQFFYACMLKYAEOLINPFGDDDDFFETMLVDR
 SIQVSLSAVDENHDLPMERDMYNDDEPHPPYTAASQSRSEFSGSTENISLKE
 DHEPPEPEBEKNTGLILGHGKFKALACRGCSYSAPOPTGLHPPVPEPSAPLGLRV
 ARLDSDSDSDSKAMREGEKSAALACRGCSYSAPOPTGLHPPVPEPSAPLGLRV
 SGIDBAANDQSLQPAVPSIKSFELPPSAEASAPLQSGSHRRRTVEFNLAADSEAP
 EHKKEPNLEPMPGIIHAIILKDRHPWALENREDAHS"
 BASE COUNT 293 a 390 c 360 g 246 t
 ORIGIN
 Query Match 26.7%; Score 648.4; DB 4; Length 1289;
 Best Local Similarity 70.3%; Pred. No. 2e-143;
 Matches 1058; Conservative 0; Mismatches 221; Indels 227; Gaps 5;
 565 CCCAGCAGCTGTGTCAGAGAGGCTTATGACTCCGACAGACAAACAGCTGGAGAAC 624
 4 CCCAGCAGCTGTGTCAGAGAGGCTTATGACTCCGACAGACAAACAGCTGGAGAAC 63
 625 TGAAGCTTACACAAACATGTTCTGGGTGCTGGGCTGTTGGCAACCTGTCAATGA 684
 64 TGAAGCTTACACAAACATGTTCTGGGTGCTGGGCTGTTGGCAACCTGTCAATGA 123
 685 AGGCGTGGCTGGAGTGCAGATCCGGGACCTTACCTGCTCCAGAGCTGTGACAGAGA 744
 124 AGGCGTGGCTGGAGTGCAGATCCGGGACCTTACCTGCTCCAGAGCTGTGACAGAGA 183
 745 TGAACACTTGCCTACTCAGTGTGACACCTGTATGCTTACGACTGATTAATGCCAC 804
 184 TGAACACTTGCCTACTCAGTGTGACACCTGTATGCTTACGACTGATTAATGCCAC 243
 805 TGGTGTATACAGAGTGTGACAGTGTGGGCTGATACACTTCTTCTGACTTGTCTAGT 864
 244 TGGTGTATACAGAGTGTGACAGTGTGGGCTGATACACTTCTTCTGACTTGTCTAGT 303
 865 GGGCGAGTTTCTGAAACCCAGAGCCTTACCTGACATGAGCTGGACCTGCTGTGTC 924
 304 GAGGCACTTCTGAAACCCAGAGCCTTACCTGACATGAGCTGGACCTGCTGTGTC 363
 925 CCGTCTTACAGTCTTCTGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 984
 364 CCGTCTTACAGTCTTCTGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 409
 985 GGGCGCTCTGAGGCTGAGGATGAGCAGAGAGGCTTATGCTCCAGAGCTGCTTGAACGA 1044


```
repeat_region 7306..7497
/rpt_family="SVA"
repeat_region 7752..7828
/rpt_family="MIR"
repeat_region 9099..9403
/rpt_family="Alu"
complement(9526..9897)
/rpt_family="MER1"
repeat_region 9635..9752
/rpt_family="MER1B"
repeat_region 11924..12197
/rpt_family="Alu"
complement(11955..12151)
/rpt_family="SVA"
repeat_region 16757..17039
/rpt_family="Alu"
complement(16775..17010)
/rpt_family="SVA"
repeat_region 18229..18485
/rpt_family="Alu"
complement(18220..18507)
/rpt_family="SVA"
repeat_region 20101..23446
/gene="HOMER1"
/note="Human ferritin heavy chain gene"
join(20101..20421,22218..22364,22621..22746,22842..23446)
/gene="HOMER1"
/product="ferritin heavy chain"
complement(21466..21731)
/rpt_family="Alu"
repeat_region 21506..21714
/rpt_family="SVA"
complement(23272..37375)
/gene="VMD2"
/note="Best macular dystrophy gene"
complement(join(23372..23629,24840..25478,27690..27841,
28155..28235,29427..29579,30274..30351,30740..30894,
31787..32020,32537..32631,35777..35964,37308..37375))
/gene="VMD2"
/product="Best macular dystrophy protein"
complement(23735..23954)
/rpt_family="Alu"
repeat_region 23789..23956
/rpt_family="SVA"
repeat_region 26425..26707
/rpt_family="Alu"
complement(26444..26696)
/rpt_family="SVA"
repeat_region 27103..27349
/rpt_family="Alu"
complement(27106..27323)
/rpt_family="SVA"
repeat_region 29161..29307
/rpt_family="Alu"
complement(29762..30043)
/rpt_family="Alu"
repeat_region 29786..30017
/rpt_family="SVA"
repeat_region 30462..30516
/rpt_family="MIR"
complement(31095..31409)
/rpt_family="Alu"
repeat_region 33005..33260
/rpt_family="Alu"
repeat_region 33247..33597
/rpt_family="LIMC2"
complement(33849..34431)
/rpt_family="Alu"
repeat_region 34076..34411
/rpt_family="SVA"
complement(34640..34890)
```

```
repeat_region /rpt_family="Alu"
34972..35401
/rpt_family="Alu"
repeat_region complement(34988..35392)
/rpt_family="SVA"
repeat_region 36238..36487
/rpt_family="Alu"
repeat_region complement(36253..36457)
/rpt_family="SVA"
repeat_region complement(38661..38855)
/rpt_family="L1"
repeat_region complement(40997..41224)
/rpt_family="Alu"
repeat_region 41056..41213
/rpt_family="SVA"
repeat_region 42745..43020
/rpt_family="Alu"
repeat_region complement(42763..42966)
/rpt_family="SVA"
repeat_region 43312..43457
/rpt_family="Alu"
repeat_region complement(43963..44217)
/rpt_family="L1"
repeat_region complement(45296..45533)
/rpt_family="SVA"
repeat_region complement(45300..45401)
/rpt_family="L1"
repeat_region complement(45407..45533)
/rpt_family="L1"
repeat_region complement(45542..45818)
/rpt_family="Alu"
repeat_region 45598..45770
/rpt_family="SVA"
repeat_region complement(45862..46004)
/rpt_family="L1"
repeat_region complement(45998..46282)
/rpt_family="Alu"
repeat_region 46022..46260
/rpt_family="SVA"
repeat_region complement(46692..47583)
/rpt_family="Alu"
repeat_region 46885..47215
/rpt_family="SVA"
repeat_region 47363..47571
/rpt_family="SVA"
repeat_region complement(47665..47773)
/rpt_family="Alu"
repeat_region complement(47792..48077)
/rpt_family="Alu"
repeat_region complement(48110..48388)
/rpt_family="L1"
repeat_region complement(48827..49102)
/rpt_family="Alu"
repeat_region 48883..49110
/rpt_family="SVA"
repeat_region complement(49395..49626)
/rpt_family="Alu"
repeat_region 50184..50437
/rpt_family="Alu"
```

Query Match 26.4%; Score 642.2; DB 9; Length 142092;
Best Local Similarity 99.5%; Pred. No. 1.3e-141; Indels 0; Gaps 0;
Matches 644; Conservative 0; Mismatches 3;

| | | | |
|----|-------|---|-------|
| QY | 1405 | CAGCTGAACAAAGAGATGATGATCCAGCCCAATCAGAGAGAGAGATGCTCA | 1464 |
| DB | 25481 | CAGCTGAACAAAGAGATGATGATCCAGCCCAATCAGAGAGAGAGATGCTCA | 25422 |
| QY | 1465 | CGCTGCAATCATTGCGCGCTTCTAGGCTGCGATCCATCAGATCCATCCAGGCG | 1524 |
| DB | 25421 | CGCTGCAATCATTGCGCGCTTCTAGGCTGCGATCCATCAGATCCATCCAGGCG | 25362 |
| QY | 1525 | AAATCAAGGACCAACATCTGTGGCCCAAGAGGGAATCCCTTCTCCAGAGGCGCTGCC | 1584 |

```

Db      25361  AAACCTAAGGACCAACTACTGTGGCCCAAGAGGGAATCCCTTCTCAGAGGCGCTGCC 25302
Qy      1555  CAAAACCAAGGACCAAGCAAGAGCTTGGGGCCAGAGAGACAAAGGCTTGAA 1644
Db      25301  CAAAACCAAGGACCAAGCAAGAGCTTGGGGCCAGAGAGACAAAGGCTTGAA 25242
Qy      1645  GCTTAAGGCTGGAGGCGCTTCAAGTGTGGGCCCACTGTATCAGAGGCGGCTACTACAG 1704
Db      25241  GCTTAAGGCTGGAGGCGCTTCAAGTGTGGGCCCACTGTATCAGAGGCGGCTACTACAG 25182
Qy      1705  TGCCCCACAGAGCGCCCTCAGGCCCACTCCATGTTCTTCCCTTGAACATCAGCGCC 1764
Db      25181  TGCCCCACAGAGCGCCCTCAGGCCCACTCCATGTTCTTCCCTTGAACATCAGCGCC 25122
Qy      1765  GTCAAGGCTTCAAGTGTACAGAGGATAGACACCAAGCAAAAGCTTAAAGCTGTGAG 1824
Db      25121  GTCAAGGCTTCAAGTGTACAGAGGATAGACACCAAGCAAAAGCTTAAAGCTGTGAG 25062
Qy      1825  TTCTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGGCCTTANTGAGCA 1884
Db      25061  TTCTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGGCCTTANTGAGCA 25002
Qy      1885  CCCAGAACTATCTCAAGTGAAGAGAAAGCTGTGAGTTTAACTGACGATATGCCAGA 1944
Db      25001  CCCAGAACTATCTCAAGTGAAGAGAAAGCTGTGAGTTTAACTGACGATATGCCAGA 24942
Qy      1945  GATCCCGGAAATATCACTCTCAAGAAACCTTTGGACATTCACCAACCAATACACACTAC 2004
Db      24941  GATCCCGGAAATATCACTCTCAAGAAACCTTTGGACATTCACCAACCAATACACACTAC 24882
Qy      2005  ACTCAAGATCACAATGATGATCTTATTTGGGCGCTTGGAAAGGAGATG 2051
Db      24881  ACTCAAGATCACAATGATGATCTTATTTGGGCGCTTGGAAAGGAGATG 24835

```

```

RESULT 6
AC087451/c 163915 bp DNA linear HTG 04-JUL-2001
LOCUS      Homo sapiens chromosome 11 clone RP11-810P12 map 11, WORKING DRAFT
SEQUENCE   7 unordered pieces.
AC087451
AC087451.2 GI:14595836
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS   Homo sapiens.
SOURCE      Homo sapiens.
ORGANISM   Homo sapiens.

```

```

REFERENCE
AUTHORS    Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.
TITLE      1 (bases 1 to 163915)
JOURNAL    Homo sapiens chromosome 11, clone RP11-810P12
REFERENCE  2 (bases 1 to 163915)
AUTHORS    Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
            Baran, N., Bastien, V., Boguski, L., Bouckge, B., Brown, A.,
            Camarata, J., Campolano, A., Choquet, Y., Colangelo, M., Collins, S.,
            Collymore, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J. S.,
            Dodge, S., Faro, S., Ferrel, P., Fitzhugh, W., Gage, D., Galagan, J.,
            Gardina, S., Glend, S., Goyette, M., Graham, L., Grand-Pierre, N.,
            Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
            Jones, C., Karatas, A., LaRocque, K., Lamazara, R., Landers, T.,
            Lebeck, J., Levine, R., Liu, G., Maclean, C., MacDonald, P.,
            Margulis, N., Matthews, C., McCarthy, M., McKean, P., McKernan, K.,
            McPherson, R., Melnick, J., Menus, L., Mihova, T., Mien, A.,
            Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H.,
            O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Patterson, K.,
            Plunkett, P., Poirier, N., Pollard, V., Raymond, C., Retter, R.,
            Roy, A., Santos, R., Schauer, S., Schupp, R., Seaman, S., Severy, P.,
            Sougnez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
            Straus, N., Subramanian, A., Talamas, J., Tefayeh, S., Theodore, J.,
            Travers, M., Travers, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
            Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zelnick, J.,

```

TITLE
JOURNAL
COMMENT

Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 4, 2001 this sequence version replaced g1:12039464.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project information

Center project name: L11863

Center clone name: 810_P_12

Sequencing vector: Plasmid; n/a; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 162458 bases at least Q40

Consensus quality: 162916 bases at least Q30

Insert size: 170000; agarose-fp

Insert size: 163315; sum-of-contigs

Quality coverage: 10.9 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 17285: contig of 17285 bp in length
* 17286 17385: gap of 100 bp
* 17386 19170: contig of 1795 bp in length
* 19171 19270: gap of 100 bp
* 19271 28600: contig of 9330 bp in length
* 28601 28700: gap of 100 bp
* 28701 47671: contig of 18971 bp in length
* 47672 47771: gap of 100 bp
* 47772 105063: contig of 57292 bp in length
* 105064 105163: gap of 100 bp
* 105164 142181: contig of 37018 bp in length
* 142182 142281: gap of 100 bp
* 142282 163915: contig of 21634 bp in length.

FEATURES

source

1. 163915

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone="RP11-810P12"

/clone_lib="RPC1-11 Human Male BAC"

1. 17285

/note="assembly-fragment"

clone_end:SP6

vector_side:left

17386. 19170

/note="assembly-fragment"

19271. 28600

/note="assembly-fragment"

28701. 47671

/note="assembly-fragment"

47772. 105063

/note="assembly-fragment"

105164. 142181

/note="assembly-fragment"

142282. 163915

/note="assembly-fragment"

/note="assembly-fragment"

| | clone_end:r7 | vector_slide:right" | | |
|------------|--------------|---------------------|-------|--------|
| BASE COUNT | 43950 | a | 37890 | c |
| ORIGIN | 43329 | t | 601 | others |

| | | | | |
|---------------------------|--------|---------------------|-----------|----------------|
| Query Match | 26.4%; | Score 642.2; | DB 2; | Length 163915; |
| Best Local Similarity | 99.5%; | Pred. NO. 1.4e-141; | | |
| Matches 644; Conservative | 0; | Mismatches 3; | Indels 0; | Gaps 0; |

| | | | |
|----|--------|---|--------|
| QY | 1405 | TAAGCTGAACAAAGAGAGATGGAGTTCAGGCCATCAAGSAGAGAGAGAGATGCTCA | 1464 |
| Db | 151430 | CAGCTGGAACAAAGAGAGATGAAGTTCCAGGCCAAATCAGAGAGAGAGAGATGCTCA | 151377 |
| QY | 1465 | CGCTGGCATATTTGGCCGCTTCTTAGGCCGTGACGTCCATGATCACCATCTCTCCAGGGC | 1524 |
| Db | 151370 | CGCTGGCATATTTGGCCGCTTCTTAGGCCGTGACGTCCATGATCACCATCTCTCCAGGGC | 151311 |
| QY | 1525 | AAATCTAAGGACCAACTACTGTGGCCCAAGAGGGAATCCCTTCTCCAGGGGCTGCC | 1584 |
| Db | 151310 | AAATCTAAGGACCAACTACTGTGGCCCAAGAGGGAATCCCTTCTCCAGGGGCTGCC | 151251 |
| QY | 1585 | CAAAAACCAAGGACGCCAAACAGAAAGCTTTAGGGGCCAGGAAGACAAACAGGCCCTGGAA | 1644 |
| Db | 151250 | CAAAAACCAAGGACGCCAAACAGAAAGCTTTAGGGGCCAGGAAGACAAACAGGCCCTGGAA | 151191 |
| QY | 1645 | GCTTAAGGCTGTGACGGCCTTCAAGTCTTGCCCACTGTATCAGAGGCCAAGGCTACTACAG | 1704 |
| Db | 151190 | GCTTAAGGCTGTGACGGCCTTCAAGTCTTGCCCACTGTATCAGAGGCCAAGGCTACTACAG | 151133 |
| QY | 1705 | TGCCCCAAGAGCGCCCTCAGGCCCACTGCCATGTTCTTCCCTTGAACCATCAGGGCC | 1764 |
| Db | 151130 | TGCCCCAAGAGCGCCCTCAGGCCCACTGCCATGTTCTTCCCTTGAACCATCAGGGCC | 151077 |
| QY | 1765 | GTCAAAAGCTTCAAGTGTACAGGACATAGACACCAAGACAAAGCAAGCTTAAAGACTGTAG | 1824 |
| Db | 151070 | GTCAAAAGCTTCAAGTGTACAGGACATAGACACCAAGACAAAGCAAGCTTAAAGACTGTAG | 151011 |
| QY | 1825 | TTTCGGGGCCCAAGAAAGTTTTGAATGCTCTCAGAGACCATGATGGGGCTTGAATGAGCA | 1884 |
| Db | 151010 | TTTCGGGGCCCAAGAAAGTTTTGAATGCTCTCAGAGACCATGATGGGGCTTGAATGAGCA | 150951 |
| QY | 1885 | CCGAGAAATATCTCAAGTAGAGAGAGAAACTGTGAAGTTTAACTACGAGATATGGCAGA | 1944 |
| Db | 150950 | CCGAGAAATATCTCAAGTAGAGAGAGAAACTGTGAAGTTTAACTACGAGATATGGCAGA | 150891 |
| QY | 1945 | GATCCCGGAAAATCACTCTCAAGAAACCTTTGGAACATCACAACCAACATACACACTAC | 2004 |
| Db | 150890 | GATCCCGGAAAATCACTCTCAAGAAACCTTTGGAACATCACAACCAACATACACACTAC | 150831 |
| QY | 2005 | ACTCAAAAGATACATGATGATCCTTATTTGGGCCCTTGAGAAACAGGAGTGC 2051 | |
| Db | 150830 | ACTCAAAAGATACATGATGATCCTTATTTGGGCCCTTGAGAAACAGGAGTGC 150784 | |

| | |
|------------|---|
| RESULT | 7 |
| AP003733 | |
| LOCUS | 166867 bp DNA linear PRI 08-DEC-2001 |
| DEFINITION | Homo sapiens genomic DNA, chromosome 11q, clone:RP11-810P12, |
| ACCESSION | AP003733 |
| VERSION | AP003733 |
| KEYWORDS | Complete sequence. |
| SOURCE | AP003733.4 GI:17426128 |
| ORGANISM | HTG. |
| | Homo sapiens DNA, clone:RP11-810P12. |
| | Homo sapiens |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. |
| REFERENCE | 1 |
| AUTHORS | Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., |
| TITLE | Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y. |
| JOURNAL | Homo sapiens genomic DNA |
| REFERENCE | Published Only In Database (2001) |
| AUTHORS | 2 (bases 1 to 166867) |
| | Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., |

| | |
|------------|--|
| TITLE | Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y. |
| JOURNAL | Direct Submission Submitted (11-JUN-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Science Center (GSC); 1-7-22 Suenho-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: hattori@psc.riken.go.jp, URL: http://hgp.psc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170) On Dec 7, 2001 this sequence version replaced gi:17026124. |
| COMMENT | |
| FEATURES | location=Ovalfliers |
| SOURCE | 1.166867 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="11" /map="11g" /clone="Rp11-810P12" |
| BASE COUNT | 44705 a 38465 c 39268 g 44429 t |
| ORIGIN | |

| | | | | |
|-----------------------|-----------------|---------------------|-----------|----------------|
| Query Match | 26.48; | Score 642.2; | DB 9; | Length 166867; |
| Best Local Similarity | 99.58; | Pred. No. 1.4e-141; | | |
| Matches 644; | Conservative 0; | Mismatches 3; | Indels 0; | Gaps 0; |

| | | | |
|----|-------|--|-------|
| QY | 1405 | CAGCCTCAACAAAGAGAGATGAGATCTCCAGCCCATTCAGAGAGACAGAGATCTCTCA | 1464 |
| Db | 12486 | CAGCCTCAACAAAGAGAGATGAGATCTCCAGCCCATTCAGAGAGACAGAGATCTCTCA | 12545 |
| QY | 1465 | CGCTGGCATATATTGGCCGCTTCTCTAGGCTGCAAGTCCCATGATCACCATCTCTCCAGGGC | 1524 |
| Db | 12346 | CGCTGGCATATATTGGCCGCTTCTCTAGGCTGCAAGTCCCATGATCACCATCTCTCCAGGGC | 12605 |
| QY | 1525 | AAACTCAAGAGCAACAACTACTGTGGCCCAAGAGGAAATCCCTTCTTCACAGAGGCGCTGCC | 1584 |
| Db | 12606 | AAACTCAAGAGCAACAACTACTGTGGCCCAAGAGGAAATCCCTTCTTCACAGAGGCGCTGCC | 12655 |
| QY | 1585 | CAAAAACCAAGAGGAGCCCAACAGAACGTTAGGGCCAGGAAGAACAAAGAGGCGTGGAA | 1644 |
| Db | 12666 | CAAAAACCAAGAGGAGCCCAACAGAACGTTAGGGCCAGGAAGAACAAAGAGGCGTGGAA | 12725 |
| QY | 1645 | GCTTAAGGCTGTGGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCCAGGCTTACTACAG | 1704 |
| Db | 12726 | GCTTAAGGCTGTGGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCCAGGCTTACTACAG | 12785 |
| QY | 1705 | TGGCCCAACAGACGCCCTCAGGCCCACTGCTCCATGTTCTTCCCTTGAACATTCAGAGCGCC | 1764 |
| Db | 12786 | TGGCCCAACAGACGCCCTCAGGCCCACTGCTCCATGTTCTTCCCTTGAACATTCAGAGCGCC | 12845 |
| QY | 1765 | GTCAAAGCTTCAACAGTGTACAGAGGCTATAGACACCAACAAAGCTTAAGACTGTGAG | 1824 |
| Db | 12846 | GTCAAAGCTTCAACAGTGTACAGAGGCTATAGACACCAACAAAGCTTAAGACTGTGAG | 12905 |
| QY | 1825 | TTCCTGGGGCCCAAGAAAGTTTGAATATGCTGCACAGACGAGTGGGGCCTTATGATGAGA | 1884 |
| Db | 12906 | TTCCTGGGGCCCAAGAAAGTTTGAATATGCTGCACAGACGAGTGGGGCCTTATGATGAGA | 12965 |
| QY | 1885 | CCCAAGAGTATCTCAAGTGAAGAGAGAGAAACTGTGAGTTTAACTGACGAGATATCCAGA | 1944 |
| Db | 12966 | CCCAAGAGTATCTCAAGTGAAGAGAGAGAAACTGTGAGTTTAACTGACGAGATATCCAGA | 13025 |
| QY | 1945 | GATCCCCGAAATCACCTCAAAAGACCTTTGGAACATTCACCAACCAACATTCACACATAC | 2004 |
| Db | 13026 | GATCCCCGAAATCACCTCAAAAGACCTTTGGAACATTCACCAACCAACATTCACACATAC | 13085 |
| QY | 2005 | ACTCAAGATCAATGGATCTTATTGGGCGCTTGGAAAAACAGGATG 2051 | |
| Db | 13086 | ACTCAAGATCAATGGATCTTATTGGGCGCTTGGAAAAACAGGATG 13132 | |

| | | | | |
|------------|--|-----|--------|-----------------|
| RESULT 8 | 196080 bp | DNA | linear | PRI 24-FEB-1999 |
| AC004228 | | | | |
| LOCUS | | | | |
| DEFINITION | Homo sapiens chromosome 11q12.2 PAC clone pDD519613 containing | | | |
| ACCESSION | human gene, for ferritin heavy chain (FTH), complete sequence. | | | |
| AC004228 | | | | |

| | | |
|-----------|---|------------|
| VERSION | AC004228.2 | GI:4263838 |
| KEYWORDS | HTG. | |
| SOURCE | Homo sapiens. | |
| ORGANISM | Homo sapiens | |
| REFERENCE | Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. | |
| AUTHORS | 1 (bases 1 to 196080) | |
| TITLE | Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., desaliboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T. | |
| JOURNAL | HTGS Submission | |
| REFERENCE | Unpublished | |
| AUTHORS | 2 (bases 1 to 196080) | |
| TITLE | Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brignac,S., Bumeister,R., Davis,C., English,C., Franklin,T.L., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harts,J., Hinson,S., Narayanaswamy,D., Newton,J., O'Brien,K., Patel,P., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R. | |
| JOURNAL | Direct Submission | |
| REFERENCE | Submitted (26-FEB-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA | |
| AUTHORS | 3 (bases 1 to 196080) | |
| TITLE | Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Butler,C., Card,P., desaliboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Waller,K. and Ward,T. | |
| JOURNAL | Direct Submission | |
| REFERENCE | Submitted (24-FEB-1999) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA | |
| AUTHORS | On Feb 24, 1999 this sequence version replaced gi:12911733. | |
| TITLE | IMPORTANT: This submission contains the entire insert of clone pJ519013. pJ519013 comes from the RPI-3 PAC library constructed at the Roswell Park Cancer Institute by the Pletier de Jong group. | |
| JOURNAL | CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11p12.2 Best's disease region mapped between STS D11S461 and EST A1NAK. This region spans over 1.5 Mbp. | |
| REFERENCE | MARKER CONFIRMATION: EST: FTH (ferritin heavy chain mRNA), STS D11S699 and WI-7524 | |
| AUTHORS | MAPPED CLONE OVERLAP: PACs pJ466a11 and pJ0756b9. | |
| TITLE | Location/Qualifiers | |
| JOURNAL | 1..196080 | |
| REFERENCE | /organism="Homo sapiens" | |
| AUTHORS | /db_xref="taxon:9606" | |
| TITLE | repeat_region complement(2210..2482) | |
| JOURNAL | /rpt_family="Alu" | |
| REFERENCE | 3457..3743 | |
| AUTHORS | /rpt_family="Alu" | |
| TITLE | repeat_region complement(20145..20758) | |
| JOURNAL | /rpt_family="Alu" | |
| REFERENCE | repeat_region complement(23194..23599) | |
| AUTHORS | /rpt_family="Alu" | |
| TITLE | repeat_region complement(22367..27685) | |
| JOURNAL | /rpt_family="Alu" | |
| REFERENCE | 28047..28457 | |
| AUTHORS | /rpt_family="Alu" | |
| TITLE | repeat_region complement(29390..29565) | |
| JOURNAL | /rpt_family="Alu" | |
| REFERENCE | repeat_region complement(31076..31255) | |
| AUTHORS | /rpt_family="Alu" | |
| TITLE | repeat_region complement(34948..35045) | |
| JOURNAL | /rpt_family="Alu" | |
| REFERENCE | 35232..35538 | |
| AUTHORS | /rpt_family="Alu" | |
| TITLE | repeat_region complement(36242..36540) | |
| JOURNAL | /rpt_family="Alu" | |

| | | |
|---------------|--------------------------|--|
| repeat_region | 37238..37522 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(38555..38650) | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(40483..41164) | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 42042..42116 | |
| repeat_region | /rpt_family="MIR" | |
| repeat_region | 42677..42817 | |
| repeat_region | /rpt_family="L1R12" | |
| repeat_region | 42856..43157 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 43148..43360 | |
| repeat_region | /rpt_family="L1R12" | |
| repeat_region | complement(43734..43999) | |
| repeat_region | /rpt_family="L1R12" | |
| repeat_region | complement(46998..47169) | |
| repeat_region | /rpt_family="MIR" | |
| repeat_region | complement(48048..48346) | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 48482..48737 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(48893..49037) | |
| repeat_region | /rpt_family="L1" | |
| repeat_region | 49118..49292 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 49791..50059 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 50436..50501 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(51147..51441) | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(52700..52979) | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 53281..53568 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 54071..54204 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 54506..54793 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 54807..55061 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(64533..64826) | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(65245..65783) | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(68465..68732) | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(70191..70476) | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 70560..70692 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 72294..72470 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(72800..73095) | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 73424..73695 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 73784..74071 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 74652..74935 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(75265..75514) | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(76086..76353) | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(78038..78312) | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 78513..78734 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | 78866..79180 | |
| repeat_region | /rpt_family="Alu" | |
| repeat_region | complement(79384..79673) | |

```

repeat_region /rpt_family="Alu" complement(81398) .81690)
repeat_region /rpt_family="Alu" 82013 .82288)
repeat_region /rpt_family="Alu" complement(83419) .83600)
repeat_region /rpt_family="MIR" complement(87320) .87624)
repeat_region /rpt_family="Alu" complement(87710) .87987)
repeat_region /rpt_family="Alu" complement(88186) .88472)
repeat_region /rpt_family="Alu" 88822 .89049)
repeat_region /rpt_family="Alu" 89224 .89308)
repeat_region /rpt_family="MER20" 89729 .90487)
repeat_region /rpt_family="Alu" 90504 .90618)
repeat_region /rpt_family="Alu" complement(93313) .93952)
repeat_region /rpt_family="Alu" 96938 .97517)
repeat_region /rpt_family="Alu" complement(104199) .104514)
repeat_region /rpt_family="Alu" complement(109738) .109979)
repeat_region /rpt_family="Alu" 110022 .110141)
repeat_region /rpt_family="Alu" complement(111091) .112016)
repeat_region /rpt_family="Alu" 113210 .113419)
repeat_region /rpt_family="L1" complement(114951) .115005)
repeat_region /rpt_family="MIR" 115685 .115791)
repeat_region /rpt_family="MIR" 116870 .117005)
repeat_region /rpt_family="MIR" complement(117877) .118172)
repeat_region /rpt_family="Alu" 118352 .118600)
repeat_region /rpt_family="Alu" 118679 .118813)
repeat_region /rpt_family="Alu" complement(121233) .121520)
repeat_region /rpt_family="Alu" 122668 .123071)

```

Query Match 26.4%; Score 642.2; DB 9; Length 196080;
 Best Local Similarity 99.5%; Pred. No. 1.4e-141;
 Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1405 CAGCGTGAACAAAGAGAGATGAGATGCTCCAGCCCAATCAGAGAGACGAGATGCTCA 1464
DB 191325 CAGCGTGAACAAAGAGAGATGAGATGCTCCAGCCCAATCAGAGAGACGAGATGCTCA 191384
QY 1465 CGGTGGCATCATGTGGCGCTTCTAGGCGCTGCAAGTCCCATGATCACCATCTCCAGGGC 1524
DB 191385 CGGTGGCATCATGTGGCGCTTCTAGGCGCTGCAAGTCCCATGATCACCATCTCCAGGGC 191444
QY 1525 AAACTGAAGGACCAAACTACTGTGGCCCAAGAGAGGAATCCCTTCTCCAGAGGGCTGCC 1584
DB 191445 AAACTGAAGGACCAAACTACTGTGGCCCAAGAGAGGAATCCCTTCTCCAGAGGGCTGCC 191504
QY 1585 CAAAACGCAAGAGGCGCAACAGACGTTAGGGGCGAGAGACAAACAAAGGCTCGAA 1644
DB 191505 CAAAACGCAAGAGGCGCAACAGACGTTAGGGGCGAGAGACAAACAAAGGCTCGAA 191564
QY 1645 GCTTAAGGCTGTGAGGCGCTTCAAGTCTGGCCCACTATCAGAGGCGCTACTACAG 1704

```

```

DB 191565 GCTTAAGGCTGTGAGGCGCTTCAAGTCTGGCCCACTATCAGAGGCGGCTACTACAG 191624
QY 1705 TGCCCCCAGAGCGCCCTCAGCCCACTCCCAATGTTTCTCCCTTAACCATACAGGGCC 1764
DB 191625 TGCCCCCAGAGCGCCCTCAGCCCACTCCCAATGTTTCTCCCTTAACCATACAGGGCC 191684
QY 1765 GTCAAAGCTTCACAGTGTACAGGATGACACCAAGACAAAAGCTTAAAGACTGTGAG 1824
DB 191685 GTCAAAGCTTCACAGTGTACAGGATGACACCAAGACAAAAGCTTAAAGACTGTGAG 191744
QY 1825 TTCTGGGGCCAGAAAAGTTTGAATTCCTCAGAGAGCGAGTGGCTTATGAGCA 1884
DB 191745 TTCTGGGGCCAGAAAAGTTTGAATTCCTCAGAGAGCGAGTGGCTTATGAGCA 191804
QY 1885 CCCAAGATATCTCAATGAGAGAGAAAACGTGTGATTTAACTGACGATATGCCAGA 1944
DB 191805 CCCAAGATATCTCAATGAGAGAGAAAACGTGTGATTTAACTGACGATATGCCAGA 191864
QY 1945 GATCCCGGAAATACCTCAAGAACCTTGGAAATCAACCAACATACACACTAC 2004
DB 191865 GATCCCGGAAATACCTCAAGAACCTTGGAAATCAACCAACATACACACTAC 191924
QY 2005 ACTCAAGATACATGATCCTTATTTGGGCTTGGAAAACAGGATG 2051
DB 191925 ACTCAAGATACATGATCCTTATTTGGGCTTGGAAAACAGGCTG 191971

RESULT 9
HSVMD2P10 706 bp DNA linear PRI 22-OCT-1998
LOCUS Homo sapiens vitelliform macular dystrophy protein (VMD2) gene,
DEFINITION exon 10
ACCESSION AF073499
VERSION AF073499.1 GI:3598873
KEYWORDS 10 of 11
SEGMENT Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Stohr,H., Marguardt,A., Rivera,A., Cooper,P.R., Nowak,N.J.,
Shows,T.B., Gerhard,D.S. and Weber,B.H.
TITLE A gene map of the Best's vitelliform macular dystrophy region in
JOURNAL chromosome 11q12-q13.1
MEDLINE Genome Res. 8 (1), 48-56 (1998)
PUBMED 9445487
REFERENCE 2 (bases 1 to 706)
AUTHORS Marguardt,A., Stohr,H., Passmore,L.A., Kramer,F., Rivera,A. and
Weber,B.H.
TITLE Mutations in a novel gene, VMD2, encoding a protein of unknown
JOURNAL properties cause juvenile-onset vitelliform macular dystrophy
MEDLINE Hum. Mol. Genet. 7 (9), 1517-1525 (1998)
PUBMED 9700209
REFERENCE 3 (bases 1 to 706)
AUTHORS Marguardt,A., Stohr,H., Passmore,L., Kraemer,F., Rivera,A. and
Weber,B.H.
TITLE Direct Submission
JOURNAL Submitted (23-JUN-1998) Human Genetics, University, Bionzentrum, Am
MEDLINE Hubland, Wuerzburg 97074, Germany
PUBMED Location/Qualifiers
FEATURES
source 1..706
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q12-q13.1"
/clone="PAC d1466a11"
36..674
/gene="VMD2"
/number=10

```

BASE COUNT 204 a 209 c 164 g 129 t

ORIGIN

Query Match 26.4%; Score 640.6; DB 9; Length 706;
 Best Local Similarity 99.4%; Pred. No. 1.3e-141;
 Matches 643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1405 CAGCCTTAACAAAGAGAGATGAGTTCAGACCCCAATCAGAGAGAGAGATGCTCA 1464
 DB 33 CAGCCTTAACAAAGAGAGATGAGTTCAGACCCCAATCAGAGAGAGAGATGCTCA 92

QY 1465 CGCGGATCATTTGGCGGCTTCCATAGGCGTCAGTCCATATCCATCTCCAGAGGC 1524
 DB 93 CGCGGATCATTTGGCGGCTTCCATAGGCGTCAGTCCATATCCATCTCCAGAGGC 152

QY 1525 AAACCTAAGACCAAACTACTGTGGCCCAAGAGGGAATCCCTTCCACGAGGCGCTCC 1584
 DB 153 AAACCTAAGACCAAACTACTGTGGCCCAAGAGGGAATCCCTTCCACGAGGCGCTCC 212

QY 1585 CAAAACCAACAGGACAGCAACAGACGTTAGGGCCGAGAGACAAACAGGCGCTGAA 1644
 DB 213 CAAAACCAACAGGACAGCAACAGACGTTAGGGCCGAGAGACAAACAGGCGCTGAA 272

QY 1645 GCTTAAAGCTGTGGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCGGCTACTAG 1704
 DB 273 GCTTAAAGCTGTGGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCGGCTACTAG 332

QY 1705 TGGCCCAACAGACGCGCCCTCAGACCCCACTCCATGTTCTTCCCTAGAACCATCAGGCC 1764
 DB 333 TGGCCCAACAGACGCGCCCTCAGACCCCACTCCATGTTCTTCCCTAGAACCATCAGGCC 392

QY 1765 GTCAAACCTTCACAGTGTACAGGATGACACCAACAAAGCTTAAAGACTGTAG 1824
 DB 393 GTCAAACCTTCACAGTGTACAGGATGACACCAACAAAGCTTAAAGACTGTAG 452

QY 1825 TTCTGGGGCCAGAGAAAGTTTGAATGCTCTCAGAGAGCGAGTGGGCGCTTGATGAGCA 1884
 DB 453 TTCTGGGGCCAGAGAAAGTTTGAATGCTCTCAGAGAGCGAGTGGGCGCTTGATGAGCA 512

QY 1885 CCCAGAGATATCTCAAGTGAAGAGAGAAAGCTGTGAGTTTAACTGACGAGATGCGAGA 1944
 DB 513 CCCAGAGATATCTCAAGTGAAGAGAGAAAGCTGTGAGTTTAACTGACGAGATGCGAGA 572

QY 1945 GATCCCCGGAATAATCACTCAAGAAACCTTTGGAAATCACCACCAACATACACATAC 2004
 DB 573 GATCCCCGGAATAATCACTCAAGAAACCTTTGGAAATCACCACCAACATACACATAC 632

QY 2005 ACTCAAGATCACATGATGATCTTPTTGGCGCTTGGAAACAGGGATG 2051
 DB 633 ACTCAAGATCACATGATGATCTTPTTGGCGCTTGGAAACAGGGATG 679

RESULT 10
 AC003025 112309 bp DNA linear PRI 23-JUL-1998
 LOCUS Human Chromosome 11p12.2 PAC clone PDJ466a11, complete sequence.
 AC003025
 VERSION AC003025.1 GI:3373308
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eulharia; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 112309)
 Evans,G.A., Athanaslou,M., Aguayo,P., Armstrong,D., Basit,M.,
 Buetner,J., Bumester,R., Card,P., desaliboat,F., Dunn,J.,
 English,C., Ehrhidge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
 Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
 Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
 Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
 HTGS Submission
 Unpublished
 2 (bases 1 to 112309)

AUTHORS Evans,G.A., Athanaslou,M., Bradbury,P., Brignac,S., Bumester,R.,
 Davis,C., English,C., Franklin,T.L., Garner,H.R., Gordon,M.,
 Gotway,G., Grant,O., Hahner,L., Harris,J., Hinson,S.,
 Narayanaswamy,U., Newton,J., O'Brien,K., Olliver,T., Patel,P.,
 Probst,S., Rayner,S., Schageman,J., Schilling,P., Schultz,R.,
 Syed,M., Valenzuela,D., Ward,T. and Wilson,R.
 Direct Submission

TITLE Submitted (21-OCT-1997) Genome Science and Technology Center,
 University of Texas Southwestern Medical Center at Dallas, 5323
 Harry Hines Blvd., Dallas, TX 75235-8591, USA
 3 (bases 1 to 112309)

REFERENCE Evans,G.A., Athanaslou,M., Aguayo,P., Armstrong,D., Basit,M.,
 Buetner,J., Bumester,R., Card,P., desaliboat,F., Dunn,J.,
 English,C., Ehrhidge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
 Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
 Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
 Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
 Direct Submission

TITLE Submitted (23-JUL-1998) Genome Science & Technology Center,
 University of Texas Southwestern Medical Center, 5323 Harry Hines
 Blvd., Dallas, TX 75235-8591, USA

COMMENT On Jul 23, 1998 this sequence version replaced gi:254967.
 IMPORTANT: PDJ466a11 comes from a PAC library constructed at the
 Roswell Park Cancer Institute by the Pieter de Jong group. This
 clone has been finished according to strict quality criteria and
 attempts have been made to resolve all base calling problems such
 as compressions and repetitive elements. The expected Phred/Phrap
 calculated errors/10kb is 0.18. In addition, this sequence has
 been finished such that 99.9% of consensus base calls consist of
 either double-stranded coverage or 2 types of labeling chemistry on
 one strand.
 Further information regarding the map of this region or
 annotation of PDJ466a11 can be found at
 http://gsstec.smed.edu/chromosome1.htm.
 CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome
 11p12.2 Best's disease region mapped between SRS D11S461 and EST
 AHNK. This region spans over 1.5 Mbp.
 MARKER CONFIRMATION: EST: FTH (ferritin heavy chain mRNA)
 MAPPED CLONE OVERLAP: HTGS submitted PAC clones PDJ519013 and
 PDJ756b9.

FEATURES
 source location/Qualifiers
 1..112309
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 complement(4341..4656)
 /rpt_family="Alu"
 complement(9880..10120)
 /rpt_family="Alu"
 10163..10282
 /rpt_family="Alu"
 complement(11232..
 /rpt_family="Alu"
 13351..13560
 /rpt_family="LI"
 complement(15092..
 /rpt_family="MIR"
 15826..15932
 /rpt_family="MIR"
 17011..17146
 /rpt_family="MIR"
 complement(18018..
 /rpt_family="Alu"
 18493..18741
 /rpt_family="Alu"
 18820..18954
 /rpt_family="Alu"
 complement(21374..
 /rpt_family="Alu"
 22809..23212
 /rpt_family="Alu"
 23499..23752
 /rpt_family="Alu"

```
repeat_region 24118..24472 /rpt_family="LFR5"
repeat_region 25005..25113 /rpt_family="LFR5"
repeat_region 31160..31442 /rpt_family="L1"
repeat_region 32651..32947 /rpt_family="L1"
repeat_region 35970..36025 /rpt_family="L1"
repeat_region 38958..39145 /rpt_family="L1"
repeat_region 44878..45480 /rpt_family="L1"
repeat_region 49769..50071 /rpt_family="L1"
repeat_region 50902..50982 /rpt_family="L1"
repeat_region 51038..51659 /rpt_family="L1"
repeat_region 54494..54769 /rpt_family="L1"
repeat_region 55529..55623 /rpt_family="L1"
repeat_region 56721..57100 /rpt_family="L1"
repeat_region 58402..58680 /rpt_family="L1"
repeat_region 59678..59904 /rpt_family="L1"
repeat_region 61149..61424 /rpt_family="L1"
repeat_region 62276..62710 /rpt_family="L1"
repeat_region 62915..63118 /rpt_family="L1"
repeat_region 63349..63455 /rpt_family="L1"
repeat_region 64339..64614 /rpt_family="L1"
repeat_region 64707..64984 /rpt_family="L1"
repeat_region 65596..65811 /rpt_family="L1"
repeat_region 66500..66753 /rpt_family="L1"
repeat_region 67311..67542 /rpt_family="L1"
repeat_region 67835..68119 /rpt_family="L1"
repeat_region 68860..69150 /rpt_family="L1"
repeat_region 69164..69272 /rpt_family="L1"
repeat_region 69354..70245 /rpt_family="L1"
repeat_region 70655..70939 /rpt_family="L1"
repeat_region 70933..71075 /rpt_family="L1"
repeat_region 71119..71395 /rpt_family="L1"
repeat_region 71404..71530 /rpt_family="L1"
```

```
repeat_region 71536..71637 /rpt_family="LFR5"
repeat_region 72720..72948 /rpt_family="L1"
repeat_region 73488..73633 /rpt_family="L1"
repeat_region 73925..74200 /rpt_family="L1"
repeat_region 75721..75955 /rpt_family="L1"
repeat_region 77185..77449 /rpt_family="L1"
repeat_region 78093..78222 /rpt_family="L1"
repeat_region 78227..78294 /rpt_family="L1"
repeat_region 80458..80707 /rpt_family="L1"
repeat_region 81544..81974 /rpt_family="L1"
repeat_region 82055..82305 /rpt_family="L1"
repeat_region 82514..83096 /rpt_family="L1"
repeat_region 83685..83940 /rpt_family="L1"
repeat_region 85336..85850 /rpt_family="L1"
repeat_region 86429..86483 /rpt_family="L1"
repeat_region 86902..87183 /rpt_family="L1"
repeat_region 87646..87792 /rpt_family="L1"
```

Query Match 26.4%; Score 640.6; DB 9; Length 112309;
Best Local Similarity 99.4%; Pred. No. 3.1e-141;
Matches 643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 1405 CAGCCTGACAAAGAGAGATGAGTTCACAGCCCAATCAGAGAGAGAGATGCTCA 1464
DB 91470 CAGCCTGACAAAGAGAGATGAGTTCACAGCCCAATCAGAGAGAGAGATGCTCA 91529
QY 1465 CGCTGGCATCTTGGCCGCTTCCAGGCTGCAATCCATCCCTCCAGGCG 1524
DB 91530 CGCTGGCATCTTGGCCGCTTCCAGGCTGCAATCCATCCCTCCAGGCG 91589
QY 1525 AAACCTAAGACCAAACTAGTGTGAGGAGGAAATCCCTTCCAGAGGCGTGGC 1584
DB 91590 AAACCTAAGACCAAACTAGTGTGAGGAGGAAATCCCTTCCAGAGGCGTGGC 91649
QY 1585 CAAAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1644
DB 91650 CAAAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 91709
QY 1645 GCTTAAGGCTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1704
DB 91710 GCTTAAGGCTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 91769
QY 1705 TGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1764
DB 91770 TGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 91829
QY 1765 GTCAAAAGCTTCAAGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1824
DB 91830 GTCAAAAGCTTCAAGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 91889
QY 1825 TTCTGGGGCCAAAGAAAGTTTGAATGCTGTGAGAGAGAGAGAGAGAGAGAGAG 1884
DB 91890 TTCTGGGGCCAAAGAAAGTTTGAATGCTGTGAGAGAGAGAGAGAGAGAGAGAG 91949
QY 1885 CCCAAGATATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1944
```

|||||

Db 91950 CCACGAGTATCTCAAGTGAAGGAGGAAACGTGTGAGTTTAACCTGACGGATATGCCAGA 92009

Qy 1945 GATCCCGGAAATACCTCAAGAACCTTTGGAACAAATCACCACCAACATACACACTAC 2004

Db 92010 GATCCCGGAAATACCTCAAGAACCTTTGGAACAAATCACCACCAACATACACACTAC 92069

Qy 2005 ACTCAAGATCAGATGATGATCTTATTTGGGCTTGGAACAGGATG 2051

Db 92070 ACTCAAGATCAGATGATGATCTTATTTGGGCTTGGAACAGGATG 92116

RESULT 11

AC084857 133683 bp DNA linear HTG 22-NOV-2000

LOCUS Homo sapiens chromosome 11 clone CTD-3231N5 map 11, WORKING DRAFT

DEFINITION SEQUENCE, 13 unordered pieces.

AC084857

AC084857.1 GI:11276215

HTG: HTGS_PHASE1; HTGS_DRAFT.

KEYWORDS Homo sapiens.

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE 1 (bases 1 to 133683)

1 (bases 1 to 133683)

Unpublished

2 (bases 1 to 133683)

Barren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Barna, N., Bastien, V., Bedalov, F., Boguslavsky, L., Boukhalil, B., Brown, A., Burkett, G., Campopiano, A., Casale, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J., Dodge, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardina, S., Ginde, S., Goyette, M., Graham, L., Grand-pierre, N., Hagos, B., Heatford, A., Horton, L., Iliev, I., Johnson, R., Jones, C., Kann, L., Karakas, A., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., MacDonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Meneses, L., Miho, T., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Sounguez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zimmer, A., and Zody, M.

Direct Submission

Submitted (22-NOV-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smilt, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

----- Project Information

Contact: sequence_submissions@genome.wi.mit.edu

Center project name: 3231_N5

Center clone name: 3231_N5

----- Summary Statistics

Sequencing vector: Plasmid: n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 126761 bases at least Q40

Consensus quality: 129952 bases at least Q30

Consensus quality: 131351 bases at least Q20

Insert size: 12800; agarose-fp

Insert size: 132483; sum-of-contigs

Quality coverage: 6.1 in Q20 bases; agarose-fp

Quality coverage: 5.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 1459: contig of 1459 bp in length

* 1460 1559: gap of 100 bp

* 1560 2520: contig of 961 bp in length

* 2521 2620: gap of 100 bp

* 2621 6334: contig of 3714 bp in length

* 6335 6434: gap of 100 bp

* 6435 11913: contig of 5479 bp in length

* 11914 12013: gap of 100 bp

* 12014 20756: contig of 8743 bp in length

* 20757 20856: gap of 100 bp

* 20857 29735: contig of 8879 bp in length

* 29736 29835: gap of 100 bp

* 29836 38645: contig of 8810 bp in length

* 38646 38745: gap of 100 bp

* 38746 52193: contig of 13448 bp in length

* 52194 52293: gap of 100 bp

* 52294 65209: contig of 12916 bp in length

* 65310 79796: contig of 14487 bp in length

* 79797 79896: gap of 100 bp

* 79897 95896: contig of 16000 bp in length

* 95897 95996: gap of 100 bp

* 95997 113419: contig of 17423 bp in length

* 113420 113519: gap of 100 bp

* 113520 133683: contig of 20164 bp in length.

FEATURES

source

1. 133683

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone="CTD-3231N5"

/clone_11b="CTD Human BAC"

1. 1459

/note="assembly_fragment"

1560..2520

/note="assembly_fragment"

2621..6334

/note="assembly_fragment"

6435..11913

/note="assembly_fragment"

12014..20756

/note="assembly_fragment"

20857..29735

/note="assembly_fragment"

29836..38645

/note="assembly_fragment"

38746..52193

/note="assembly_fragment"

52294..65209

/note="assembly_fragment"

65310..79796

/note="assembly_fragment"

79897..95896

/note="assembly_fragment"

95997..113419

/note="assembly_fragment"

113520..133683

/note="assembly_fragment"

BASE COUNT 31520 a 34761 c 35336 g 30857 t 1209 others

ORIGIN

Query Match 26.4%; Score 640.6; DB 2; Length 133683;

| FEATURES | * 142130 | 160169: contig of 18040 bp in length. |
|---------------------------|--|---------------------------------------|
| source | Location/Qualifiers | |
| | 1. 160169 | |
| | /organism="Homo sapiens" | |
| | /db_xref="taxon:9606" | |
| | /chromosome="11" | |
| | /clone="Rp11-810P12" | |
| | 1. 1926 | |
| misc_feature | /note="assembly_name:Contig18" | |
| | 2027. 4202 | |
| misc_feature | /note="assembly_name:Contig19" | |
| | 4303. 7163 | |
| misc_feature | /note="assembly_name:Contig20 | |
| | clone_end:17 | |
| | vector_side:left" | |
| | 7264. 9961 | |
| misc_feature | /note="assembly_name:Contig21" | |
| | 10062. 12575 | |
| misc_feature | /note="assembly_name:Contig22" | |
| | 12676. 15545 | |
| misc_feature | /note="assembly_name:Contig23" | |
| | 15646. 20941 | |
| misc_feature | /note="assembly_name:Contig24" | |
| | 21042. 23956 | |
| misc_feature | /note="assembly_name:Contig25" | |
| | 24067. 27231 | |
| misc_feature | /note="assembly_name:Contig26" | |
| | 27332. 30550 | |
| misc_feature | /note="assembly_name:Contig27" | |
| | 30651. 35202 | |
| misc_feature | /note="assembly_name:Contig28" | |
| | 35303. 40201 | |
| misc_feature | /note="assembly_name:Contig29" | |
| | 40302. 45758 | |
| misc_feature | /note="assembly_name:Contig30" | |
| | 45859. 51249 | |
| misc_feature | /note="assembly_name:Contig31" | |
| | 51350. 58267 | |
| misc_feature | /note="assembly_name:Contig32" | |
| | 58368. 66916 | |
| misc_feature | /note="assembly_name:Contig33" | |
| | 67017. 77145 | |
| misc_feature | /note="assembly_name:Contig34" | |
| | 77246. 86493 | |
| misc_feature | /note="assembly_name:Contig35" | |
| | 86594. 96688 | |
| misc_feature | /note="assembly_name:Contig36" | |
| | 96789. 107169 | |
| misc_feature | /note="assembly_name:Contig37" | |
| | 107270. 117047 | |
| misc_feature | /note="assembly_name:Contig38" | |
| | 117148. 127120 | |
| misc_feature | /note="assembly_name:Contig39 | |
| | clone_end:SP6 | |
| | vector_side:right" | |
| | 127221. 142029 | |
| misc_feature | /note="assembly_name:Contig40" | |
| | 142130. 160169 | |
| misc_feature | /note="assembly_name:Contig41" | |
| | 160169. 2307 | |
| BASE COUNT | 41647 a 36693 c 37296 g 42226 t 2307 others | |
| ORIGIN | | |
| Query Match | 22.8%; Score 552.8; DB 2; Length 160169; | |
| Best Local Similarity | 99.3%; Pred. No. 2.5e-120; | |
| Matches 576; Conservative | 0; Mismatches 2; Indels 2; Gaps 2; | |
| DB | 1405 CAGCCTGAACAAGAGAGATGAGATTCACAGCCCAATCAGAGAGACGAGAGAGATGCTCA 1464 | |
| | | |
| DB | 67594 CAGCCTGAACAAGAGAGATGAGATTCACAGCCCAATCAGAGAGACGAGAGAGATG-TC 67556 | |
| QY | 1465 CGCTGACATCATTTGGCCGCTTCCTAGGGCTGAGGCCATGATCAGCATGCTCCAGGGC 1524 | |
| | | |
| QY | 1465 CGCTGACATCATTTGGCCGCTTCCTAGGGCTGAGGCCATGATCAGCATGCTCCAGGGC 67476 | |
| | | |
| DB | 67535 CGCTGACATCATTTGGCCGCTTCCTAGGGCTGAGGCCATGATCAGCATGCTCCAGGGC 67476 | |

| QY | 1525 | AAACTCAGAGCAACAACTACTGTGGCCCAAGAGGAATCCCTTCTCCACGAGGGCCTGCC | 1584 |
|------------|--|--|-----------------|
| Db | 67475 | AAACTCAGAGCAACAACTACTGTGGCCCAAGAGGAATCCCTTCTCCACGAGGGCCTGCC | 67417 |
| QY | 1585 | CAAAAACACAGGCGACCCAAAGAAAGCTTAGGGCCAGGAGACAAACAAGGCTGTGAA | 1644 |
| Db | 67416 | CAAAAACACAGGCGACCCAAAGAAAGCTTAGGGCCAGGAGACAAACAAGGCTGTGAA | 67355 |
| QY | 1645 | GCTTAAGGCTGTGGAGCCTTCAAGTGTGGCCCACTGTATCAGAGGCGCAGCTACTACAG | 1704 |
| Db | 67356 | GCTTAAGGCTGTGGAGCCTTCAAGTGTGGCCCACTGTATCAGAGGCGCAGCTACTACAG | 67297 |
| QY | 1705 | TGCCCCACAGAGCGCCCTCAGCGCCCACTCCCATGTCTTCCGCCCTAGAACCATCAGCGCC | 1764 |
| Db | 67296 | TGCCCCACAGAGCGCCCTCAGCGCCCACTCCCATGTCTTCCGCCCTAGAACCATCAGCGCC | 67237 |
| QY | 1765 | GTCAACAGCTTCAGAGTGTACAGGCAAGCATACACCAAGAAAGAAAGCTTAAAGACTGTGAG | 1824 |
| Db | 67236 | GTCAACAGCTTCAGAGTGTACAGGCAAGCATACACCAAGAAAGAAAGCTTAAAGACTGTGAG | 67177 |
| QY | 1825 | TTCTGGGGCCCAAGAAAGTTTGAATTTGCTCTCAGAGAGCGATGGGGCCTTGATGAGCA | 1884 |
| Db | 67176 | TTCTGGGGCCCAAGAAAGTTTGAATTTGCTCTCAGAGAGCGATGGGGCCTTGATGAGCA | 67111 |
| QY | 1885 | CCCAAGATATCTCAAGTGTAGAGAGAGAAACCTGTGTAGTTTAACCTGAGCGGATATGCCAGA | 1944 |
| Db | 67116 | CCCAAGATATCTCAAGTGTAGAGAGAGAAACCTGTGTAGTTTAACCTGAGCGGATATGCCAGA | 67055 |
| QY | 1945 | GATCCCGGAAATCAGCTCAAGAAACCTTTGGAACAATCA | 1984 |
| Db | 67056 | GATCCCGGAAATCAGCTCAAGAAACCTTTGGAACAATCA | 67017 |
| RESULT 13 | | | |
| AF440756 | | 1908 bp | mRNA |
| LOCUS | | 1908 bp | linear |
| DEFINITION | | Human sapiens vitelliniform macular dystrophy 2-like protein 1 mRNA, | PRI 12-JUL-2002 |
| ACCESSION | AF440756 | | |
| VERSION | AF440756 | | |
| KEYWORDS | AF440756.1 GI:21734839 | | |
| SOURCE | human. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | |
| TITLE | 1 (bases 1 to 1908) | | |
| REFERENCE | Marquardt, A., Stohr, H., Passmore, L.A., Kramer, F., Rivera, A. and | | |
| AUTHORS | Weber, B.H. | | |
| TITLE | Mutations in a novel gene, VMD2, encoding a protein of unknown | | |
| REFERENCE | properties cause juvenile-onset vitelliniform macular dystrophy | | |
| AUTHORS | (Best's disease) | | |
| TITLE | Hum. Mol. Genet. 7 (9), 1517-1525 (1998) | | |
| JOURNAL | 98367043 | | |
| MEDLINE | 9700209 | | |
| PUBMED | 2 (bases 1 to 1908) | | |
| REFERENCE | Stohr, H., Marquardt, A., Nanda, I., Schmid, M. and Weber, B.H. | | |
| AUTHORS | Three novel human VMD2-like genes are members of the evolutionary | | |
| TITLE | highly conserved RFP-TM family | | |
| REFERENCE | Eur. J. Hum. Genet. 10 (4), 281-284 (2002) | | |
| JOURNAL | 22027749 | | |
| MEDLINE | 12032738 | | |
| PUBMED | 3 (bases 1 to 1908) | | |
| REFERENCE | Stohr, H., Marquardt, A. and Weber, B.H.F. | | |
| AUTHORS | Three novel human VMD2-like protein genes are members of the | | |
| TITLE | evolutionary highly conserved RFP family | | |
| JOURNAL | Unpublished | | |
| MEDLINE | 4 (bases 1 to 1908) | | |
| PUBMED | Stohr, H., Marquardt, A. and Weber, B.H.F. | | |
| REFERENCE | Direct Submission | | |
| AUTHORS | Submitted (30-OCT-2001) Human Genetics, University of Wuerzburg, | | |
| TITLE | Biozentrum, Am Hubland, Wuerzburg 97074, Germany | | |
| JOURNAL | Location/Qualifiers | | |

source
1. 1908
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.12-p13.2"
1.1530
/note="VMDL1: putative membrane protein; Pfam01062;
worm_family:8;
worm_start-1
/product="Vitellogenin macular dystrophy 2-like protein 1"
/protein_id="AA076995.1"
/db_xref="GI:21734840"
/translation="MTVTAVANARFGSGSOLLILMRGSIYLKMLBELCTGFTY
ALSAAYRPLTSGKRFYKIVTICDOYASLPVSLFGLVTLVNVNWSQICMPL
PALMCVAVAGTHGDDRRRLYRPTMAGISAVLILRSVTAFAKRPIDHVEA
GPMREERKFEENLSSINKYPCWMSNLAAORRIRLIDNSALKLELLELVEF
GKGMFLHYDMISVPLVYQVNTALYSTFLKGLRGROFLDPAOCYKNDHDLICVIF
TLIOFFYHGMKVAEQLINPGEDEDDDEFLLIDRNFQVSLAVDEMDYDLAVLEK
DIYMDAERAPYATVATVQLRPSFGSGTFTLAKEDMQRLDGLDGPMEAPGD
PIORLIPAGAGVAGPLGRRLSFLRKNSVSEASTKSCSAVVPBGAPECCSD
PLDPLPEPAPPPAPPPPLILIPGPVPSIVMPGPGAPWLPSPIGEEBENT
A"

BASE COUNT 337 a 604 c 551 g 416 t

ORIGIN

Query Match 17.7%; Score 431; DB 9; Length 1908;
Best Local Similarity 68.4%; Pred. No. 1.1e-21;
Matches 596; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

105 ATGACATCACTACACAGCCAGAGTGGCTTAATGCCGCTTACGCTCCCGCTG 164
1 ATGACGCTACACACCCGAGTGGCCAGACCCGCTCGGCTCTCCAGCTG 60
165 CTGCTGCTGCGGGGAGCATCTACAGCTGCTATATAGCGAGTCTTAACTCTG 224
61 CTGCTACTGCTGGCGAGCATCTACAACTCTGCGGAGAGTCTGCTGCTCT 120
225 CTGCTACTACATCACTCCGCTTATTAATAGCTGGCCCTCAGGAGAAACACAGCTG 284
121 GGGTCTACATGCGGCTGAGTGTGCTGACCGCTTGTGCTGACGAGGAGGAGGAGCGC 180
285 ATGTTGGAACAGTACCTGTATTTGAGAGAGTACATCCAGCTATCCCATTTCTTC 344
181 TACTTCGAGAGCTGTATTTATTTGTAACAGTATGCCAGCTATCCGCTCTCC 240
345 GTGCTGGCTCTACAGTACGCTGTGTGACCCGCTGTGTGAGAACCACTACGAGAACTG 404
241 GTGCTGGCTTATGTAAGCTGTGTGTAACCGCTGTGTGAGCACTATCTATGATG 300
405 CCGTGGCCGAGCGGCTCATGAGCTGTGTGCGGCTTCGTCGAGAGGAGGAGGAGCA 464
301 CCGTGGCCGAGCGGCTCATGAGCTGTGTGCGGCTTCGTCGAGAGGAGGAGCA 360
465 GGGCGGCTCTGCGGCGGAGCGCTACGCTACGCAACCTGGGAGCGTCTCATCTG 524
361 GGGCGGCTCTACGCGGAGCACTCATGCTACGCGGCTCTCGGCGGCTCATCTG 420
525 CCGAGGCTACGACCGCAGTCTACAGCGCTTCCCGAGCGCCAGCACTGCTGCAAGA 584
421 CCGTGGCTACGAGCGGCTGTCAAGCGCTTCCCGACCATGACCACTGCTGAGAGGCT 480
585 GGGCTTATGATCCGCGAGCAACAGAGTGGAGAACTGAGCTTACCAACACATG 644
481 GGGCTTATGATCCGCGAGGAGCGCAAGAGTTGAAACCTGAACTCATCTACACAG 540
645 TTGCGGTGCTGCGGTGTGTGCAACCTGTCAATGAGCGGTGCTGAGAGTCA 704
541 TACTGGTGGCTGCGGTGTGTGCAACCTGCGGAGCAAGCGCCGAGCGGCGCG 600
705 ATCCGGAGACCTATCTGCTCAGAGAGCTGTGAGAGAGTAAACCTTCTGCTACTG 764
601 ATCCGGAGACACCGCGCTTAAGCTGTGCTGAGAGTGAATGTTTTCGGGCGCAA 660

765 TGTGACACCTGATATGCTACAGCTGATGATATCCACCTGCTATACAGAGTGTG 824
661 TGTGAGATGCTCTTCTACATGACAGGATGATACACCTGCTACACGAGGTGTG 720
825 ACTGTGGCGGTATACAGCTTCTTCTGACTTGTCTAGTGTGGGCGGAGTCTGAACCA 884
721 ACCATGCACTGTACAGCTACTCTGCTGCTCCATGATGTCGCCAGTCTCTGAGACCG 780
885 GCCAAGGCTTACCTGCGGAGTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 944
781 GCTCAGGCTTACCAAGACACGACCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 840
945 TTCTCTTCTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 975
841 TTCTCTTCTACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 871

Db

RESULT 14
BC031186
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

BC031186 1956 bp mRNA linear ROD 07-AUG-2002
Mus musculus, clone MGC:37621 IMAGE:4989959, mRNA, complete cds.
BC031186 GI:21411099
MGC.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1956)
Strausberg, R.
Direct Submission
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) medpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 58 Row: 9 Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analyses.

FEATURES
source
location/Qualifiers
1. 1956
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:37621 IMAGE:4989959"
/tissue_type="Colon, normal, 5 month old male mouse."
/clone_id="NCL_CGAP_C024"
/lab_host="DH10B"
/note="vector: pcMV-SPORT6"
176..1573
/codon_start=1
/product="Unknown (protein for MGC:37621)"
/protein_id="AAH31186.1"
/db_xref="GI:21411100"
/db_xref="LocusID:212989"
/translation="MALSAAYRPLTSGKRFYKIVTICDOYASLPVSLFGLVTLVNVNWSQICMPL
PALMCVAVAGTHGDDRRRLYRPTMAGISAVLILRSVTAFAKRPIDHVEA
GPMREERKFEENLSSINKYPCWMSNLAAORRIRLIDNSALKLELLELVEF
GKGMFLHYDMISVPLVYQVNTALYSTFLKGLRGROFLDPAOCYKNDHDLICVIF
TLIOFFYHGMKVAEQLINPGEDEDDDEFLLIDRNFQVSLAVDEMDYDLAVLEK
DIYMDAERAPYATVATVQLRPSFGSGTFTLAKEDMQRLDGLDGPMEAPGD
PIORLIPAGAGVAGPLGRRLSFLRKNSVSEASTKSCSAVVPBGAPECCSD
PLDPLPEPAPPPAPPPPLILIPGPVPSIVMPGPGAPWLPSPIGEEBENT
A"

| | | | | | |
|-----------------------|---|-------------------|---|----------------|----------|
| BASE COUNT | 411 a | 577 c | 530 g | 438 t | |
| ORIGIN | | | | | |
| Query Match | 17.1% | Score 415 | DB 10 | Length 1956 | |
| Best Local Similarity | 67.0% | Pred. No. 6.9e-88 | | | |
| Matches | 589 | Conservative | 0 | Mismatches 290 | Indels 0 |
| | | | | Gaps 0 | |
| 97 | GCCTGGCCATACATCACTTACACAGCCAAAGTGCCTTAAGCCCTTACGCTCTCT | 156 | NSALKLLELVNFRSKCMLEHYMISIPYVTVYVAVSYFLACILGRFLDPA | | |
| 99 | GCGCAGCATGACCGTACCTACACAGCCAAAGTGCCTTAAGCCCTTACGCTCTCT | 158 | OQYKDHLDLCVPIFTLLOFFELAGLWAKAEOLINPGEEDDDDEFENFLIDNFOVM | | |
| 157 | CCCGCTGCTGCTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 216 | LAVDEYDLDLAMEKDLVMDAARAPYATATFLLOQSPGSPEDTALAKEDFO | | |
| 159 | CCGACCTGCTGCTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 218 | RLDVGDPLEGVHGFLORLPLPAGAGSVPLGRSLRLKRSKCSSEASTAKSCCAG | | |
| 217 | TCTTCTGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 276 | AADGGVGGCGDPLDPSLRPELEPPACPEPPAPIPGPTEPFTTVSIPGPAPAP | | |
| 219 | AGAACGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 278 | PMLPSPIGEEESPAP | | |
| 277 | AGAACGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 336 | | | |
| 279 | TCTTCTGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 338 | | | |
| 337 | TTTCTTCTGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 396 | | | |
| 339 | TATGATGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 398 | | | |
| 397 | AGAACGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 456 | | | |
| 457 | AGCAGCAAGCCGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 516 | | | |
| 459 | TGATCTTCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 518 | | | |
| 517 | TCATCTGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 576 | | | |
| 519 | TCGAGCTGCTGTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 578 | | | |
| 577 | TCGAGCTGCTGTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 636 | | | |
| 579 | ACACATGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 638 | | | |
| 637 | ACACATGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 696 | | | |
| 639 | AGGCGGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 698 | | | |
| 697 | AGGCGGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 756 | | | |
| 699 | GGAGCAAGTGTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 758 | | | |
| 817 | AGGCGGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 876 | | | |
| 819 | AGGCGGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 878 | | | |
| 877 | AGGCGGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 936 | | | |
| 879 | AGGCGGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 938 | | | |
| 937 | AGGCGGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 996 | | | |
| 939 | AGGCGGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 998 | | | |
| 997 | AGGCGGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 1056 | | | |
| 999 | AGGCGGCTGCTGTGTGTGGGGGAGCATCTACAAAGCTGCTATATGGCGAGTCTTAA | 1058 | | | |

| | | | | |
|--|---------|------|--------|-----------------|
| RESULT 15 | 1957 bp | mRNA | linear | ROD 07-AUG-2002 |
| BC019528 | | | | |
| LOCUS | | | | |
| DEFINITION | | | | |
| ACCESSION | | | | |
| VERSION | | | | |
| KEYWORDS | | | | |
| SOURCE | | | | |
| ORGANISM | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| REMARK | | | | |
| COMMENT | | | | |
| Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov | | | | |
| Series: IRK Plate: 37 Row: h Column: 19 | | | | |
| This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein. | | | | |
| location/Qualifiers | | | | |
| 1. 1957 | | | | |
| /organism="Mus musculus" | | | | |
| /db_xref="taxon:10090" | | | | |
| /map="FVB/N" | | | | |
| /clone="MGC:28568 IMAGE:4208860" | | | | |
| /tissue_type="Colon, normal, 5 month old male mouse." | | | | |
| /clone_id="MCI_CGAP_Co24" | | | | |
| /lab_host="DH10B" | | | | |
| /note="vector: pCMV-Sport6" | | | | |
| 189. 1586 | | | | |
| /codon_start=1 | | | | |
| /product="Similar to hypothetical protein FLJ20132" | | | | |
| /protein_id="AA19528.1" | | | | |
| /db_xref="GI:18044531" | | | | |
| /translation="MALSAAYRFLAIEOKRYFEKLVITYCDQYASLIVSVLDFPYT /LYHRMNOYLCPDLPALMCIVAGTVHSDRGRILYRTLMRAGLSAVILHSVT AVFRKPTLIHYVAGMTRFEERKFNENSSINRYWPCVWFSYSLAQAQREBRID NSALKLLELVNFRSKCMLEHYMDISINPLYVTOVYIAVSYFLACILGRFLDPA RAGYDHTLDLCVPIFTLLOFFELAGLWAKAEOLINPGEEDDDDEFENFLIDNFO VM LAYDEMDDLAMEKDLVMDAARAPYATATFLLOQSPGSPEDTALAKEMOFCR RLVDVGDPLEGVHGFLORLPLPAGAGSVPLGRSLRLKRSKCSSEASTAKSCCAG AADGGVGGCGDPLDPSLRPELEPPACPEPPAPIPGPTEPFTTVSIPGPAPAP PMLPSPIGEEESPAP" | | | | |

| | | | | | |
|-----------------------|--|-------------------|---|----------------|----------|
| BASE COUNT | 402 a | 584 c | 534 g | 437 t | |
| ORIGIN | | | | | |
| Query Match | 17.1% | Score 415 | DB 10 | Length 1957 | |
| Best Local Similarity | 67.0% | Pred. No. 6.9e-88 | | | |
| Matches | 589 | Conservative | 0 | Mismatches 290 | Indels 0 |
| | | | | Gaps 0 | |
| 97 | GCCTGGCCATACATCACTTACACAGCCAAAGTGCCTTAAGCCCTTACGCTCTCT | 156 | NSALKLLELVNFRSKCMLEHYMISIPYVTVYVAVSYFLACILGRFLDPA | | |

RESULT 15

BC019528 1957 bp. mRNA linear. ROD 07-AUG-2002

LOCUS

Mus musculus, similar to hypothetical protein FLJ20132, clone

DEFINITION

MGC:28568 IMAGE:4208860, mRNA, complete cds.

ACCESSION

BC019528

VERSION

BC019528.1 GI:18044530

KEYWORDS

MGC.

SOURCE

house mouse.

ORGANISM

Mus musculus.

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLES

1 (bases 1 to 1957)

JOURNAL

Strausberg, R.

REMARK

Direct Submission

COMMENT

Submitted (19-DEC-2001) National Institutes of Health, Mammalian

FEATURES

Gene Collection (MGC), Cancer Genomics Office, National Cancer

SOURCE

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

FEATURES

USA

COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

FEATURES

Contact: MGC help desk

COMMENT

Email: cyabs-remail.nih.gov

FEATURES

Tissue Procurement: Jeffrey E. Green, M.D.

COMMENT

cDNA Library Preparation: Life Technologies, Inc.

FEATURES

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

COMMENT

DNA sequencing by: Baylor College of Medicine Human Genome

FEATURES

Sequencing Center:

COMMENT

Center code: BCM-HGSC

FEATURES

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

COMMENT

Contact: angbcm.tmc.edu

FEATURES

Guaratine, P.H., Garcia, A.M., Lu, X., Huliy, S.W., Hale, S.M.,

COMMENT

Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,

FEATURES

Richards, S., Gibbs, R.A.

COMMENT

Clone distribution: MGC clone distribution information can be found

FEATURES

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

COMMENT

Series: IRAC Plate: 37 Row: h Column: 19

FEATURES

This clone was selected for full length sequencing because it

COMMENT

passed the following selection criteria: similarity but not

FEATURES

identity to protein.

COMMENT

Location/Qualifiers

FEATURES

1. 1957

COMMENT

/organism="Mus musculus"

FEATURES

/db_xref="taxon:10090"

COMMENT

/map="FVB/N"

FEATURES

/clone="MGC:28568 IMAGE:4208860"

COMMENT

/tissue_type="Colon, normal. 5 month old male mouse."

FEATURES

/clone_id="NCLCGAP_Co24"

COMMENT

/lab_host="DH10B"

FEATURES

/note="Vector: pCMV-SPORT6"

COMMENT

189. 1586

FEATURES

/codon_start=1

COMMENT

/product="Similar to hypothetical protein FLJ20132"

FEATURES

/protein_id="AAH19528.1"

COMMENT

/db_xref="GI:18044531"

FEATURES

/translation="MAISAATRELLAEORKEFEKLYIVCDYASLIPSEFVLEFYVT

COMMENT

/translat="MAISAATRELLAEORKEFEKLYIVCDYASLIPSEFVLEFYVT

FEATURES

LAVHRMNOYLCLMPLDMLMCIIVAGTVHGRDRDLRLRLMAYAGLSAVLLISVST

COMMENT

AVKRFPTIDVAVGEMTREBRKFEKFNLSYKRVPCVWSSLAQAORRGRIED

FEATURES

NSALKLLELVNFRSKCMLEHYMISIPYVTVYVAVSYFLACILGRFLDPA

COMMENT

OQYKDHLDLCVPIFTLLOFFELAGLWAKAEOLINPGEEDDDDEFENFLIDNFOVM

FEATURES

LAVDEYDLDLAMEKDLVMDAARAPYATATFLLOQSPGSPEDTALAKEDFO

COMMENT

RLDVGDPLEGVHGFLORLPLPAGAGSVPLGRSLRLKRSKCSSEASTAKSCCAG

FEATURES

AADGGVGGCGDPLDPSLRPELEPPACPEPPAPIPGPTEPFTTVSIPGPAPAP

COMMENT

PMLPSPIGEEESPAP

```
Db      52  GCGAGCGATGACCGTCACTACACAGCCAGAGTGGGAAATGCCCGCTCGTGGCTTC 111
QY      157  CCGCGCTGCTGCTGCTGCGGGGACACATCTACAAAGCTCTATATGGGAGTTCTTAA 216
Db      112  CGCAGCTGCTGCTGCTGCGGGGACATCTACAAAGCTCTGCTGGGAGAGCTGTAT 171
QY      217  TCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276
Db      172  GTTTCCTGGGACTTACATGACCTAAGCGCCCTATCGCTTCTACTGGCAGAAAGAGC 231
QY      277  AACAGCTGATGTTTGAACACTGACTGTATATGCGACAGCTACATCCAGCTCATCCCA 336
Db      232  AGAAGCGCTACTTCGAGAAAGCTTGTATATACGTGACACAGTACGCGAGCTCATCCCG 291
QY      337  TTTCTTCTGCTGCTGCTTCTACGTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396
Db      292  TCTTCTGCTGCTGCTGCTTCTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 351
QY      397  AGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 456
Db      352  TATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 411
QY      457  ACGAGCAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 516
Db      412  ACGATCGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 471
QY      517  TCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 576
Db      472  TGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 531
QY      577  TGCAGAGGCTTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 636
Db      532  TCGAGGCTGATTTATGACCGGAGAGCGGCAAGAGTTGAGAACTTGAAATCGCTTC 591
QY      637  ACAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
Db      592  ACAACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
QY      697  GAGGTGATCCGGGACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756
Db      652  AAGGGGCGCATCCGCGACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 711
QY      757  GTACTGAGTGTGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816
Db      712  GGAGCAATGTGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 771
QY      817  AGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 876
Db      772  AGGTAGTCACTATCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 831
QY      877  TGAACCGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 936
Db      832  TAGACCTGCGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
QY      937  TCTGCAATCTTCTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 975
Db      892  TGTGCAATCTTCTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
```

Search completed: June 8, 2003, 18:10:36
Job time : 6129.06 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 10:51:03 ; Search time 3214.69 Seconds

(without alignments)
12237.209 Million cell updates/sec

Title: US-09-622-964-4

Perfect score: 2429
Sequence: 1 caggagagtcacacagccta.....aaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_estr:*
9: gb_estl:*
10: gb_estl2:*
11: gb_estc3:*
12: gb_estc4:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1587 | 65.3 | 2453 | 11 | BC015220 Homo sapi |
| 2 | 640.4 | 26.4 | 666 | 13 | BI480798 HZRP-043 |
| 3 | 582.4 | 24.0 | 592 | 13 | BM663028 UI-E-C10- |
| 4 | 577 | 23.8 | 773 | 10 | BE410951 601303662 |
| 5 | 575 | 23.7 | 585 | 14 | BM707948 UI-E-C11- |
| 6 | 565.2 | 23.3 | 593 | 10 | BE385296 601277572 |

| Result | LOCUS | DEFINITION | ACCESSION | VERSION | KEYWORDS | ORGANISM | REFERENCE | AUTHORS | TITLE | JOURNAL |
|--------|----------|--|-----------|---------|-------------|--------------|-----------|----------------|--|---------|
| 1 | BC015220 | Homo sapiens, clone IMAGE:3877806, mRNA. | BC015220 | 1 | GI:21955361 | Homo sapiens | 1 | Strausberg, R. | Submitted (01-Oct-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | |
| 2 | BC015220 | Homo sapiens, clone IMAGE:3877806, mRNA. | BC015220 | 1 | GI:21955361 | Homo sapiens | 2 | Strausberg, R. | Submitted (01-Oct-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | |
| 3 | BC015220 | Homo sapiens, clone IMAGE:3877806, mRNA. | BC015220 | 1 | GI:21955361 | Homo sapiens | 3 | Strausberg, R. | Submitted (01-Oct-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | |
| 4 | BC015220 | Homo sapiens, clone IMAGE:3877806, mRNA. | BC015220 | 1 | GI:21955361 | Homo sapiens | 4 | Strausberg, R. | Submitted (01-Oct-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | |
| 5 | BC015220 | Homo sapiens, clone IMAGE:3877806, mRNA. | BC015220 | 1 | GI:21955361 | Homo sapiens | 5 | Strausberg, R. | Submitted (01-Oct-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | |

ALIGNMENTS

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: WGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DP/Genzdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mdc@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://limage.llnl.gov>
Series: IRK Plate: 14 Row: b Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis

This clone has the following problem: incomplete processing.

FEATURES

source

Location/Qualifiers
1. .2453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3877806"
/issue_type="lung_large cell carcinoma"
/clone_lib="NIH_MGC_68"
/lab_host="DH10B"
/note="Vector: PCMV-Sport6"

BASE COUNT

633 a 711 c 606 g 503 t

ORIGIN

Query Match

65.38; Score 1587; DB 11; Length 2453;

Best Local Similarity 86.7%; Pred. No. 3.6e-224; Index 284; Gaps 1;

Matches 1884; Conservative 0; Mismatches 5;

257 GCTGGCCCTCAGGAGAGACACAGCTGATGTTGAGAACTGACTGTATTCGACAG 316
|||||
557 GCTGGCCCTCAGGAGAGACACAGCTGATGTTGAGAACTGACTGTATTCGACAG 616
|||||
317 CTACATCCAGCTCATCCCATTTCTCTGCTGCTGCTTCTAGCTGACGCTGCTGAC 376
|||||
617 CTACATCCAGCTCATCCCATTTCTCTGCTGCTGCTTCTAGCTGACGCTGCTGAC 676
|||||
377 CCGCTGGTGAACAGATACGAGACCTGCGCCGACCGCTCATGAGCCTGCTGTC 436
|||||
677 CCGCTGGTGAACAGATACGAGACCTGCGCCGACCGCTCATGAGCCTGCTGTC 736
|||||
437 GGGCTTCGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 496
|||||
737 GGGCTTCGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 796
|||||
497 GCGCAACCTGGGCAACGCTGCTATCTGCGAGCGCTGACGCGAGCTACAGAGCTT 556
|||||
797 GCGCAACCTGGGCAACGCTGCTATCTGCGAGCGCTGACGCGAGCTACAGAGCTT 856
|||||
857 CCGCAGCGCCAGCACTGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 916
|||||
617 GGAGAACTGAGCTTACACACACACATGTTCTGGGTGCGCTGGGTGGTTGGCAACT 676
|||||
917 GGAGAACTGAGCTTACACACACATGTTCTGGGTGCGCTGGGTGGTTGGCAACT 976
|||||
677 GTCAAG 736
|||||
977 GTCAAG 1036
|||||
737 GAAGAGATGAGACACCTTGCATGCTAGTGTGAGACCTGCTATGCTTACAGCTGAT 796
|||||
1037 GAAGAGATGAGACACCTTGCATGCTAGTGTGAGACCTGCTATGCTTACAGCTGAT 1096
|||||
797 TATCCCACTGCTATATACAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 856
|||||
1097 TATCCCACTGCTATATACAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1156
|||||
857 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTACCTGGCCATGAGCTGACCT 916
|||||
1157 TCTAGTGGGCGGAGTTTCTGAACCCAGCCAGGCTACCTGGCCATGAGCTGACCT 1216
|||||
917 CGTTGTCGCCGCTTCACTGCTTCTGAGCTTCTTATGTTGGCTGGCTGAAGTGGG 976
|||||
1217 CGTTGTCGCCGCTTCACTGCTTCTGAGCTTCTTATGTTGGCTGGCTGAAGTGGG 1269
|||||

977 CTTCTCCAGGAGCCCTGCTGGGCTGAGAGCATGGCCAGAGGGGTATGCGCAGCAGCTGCT 1036
----- 1269
1270 ----- 1269
1037 TGAGACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1096
----- 1269
1270 ----- 1269
1097 GTGGCGACACCTGATATCCAGCTACTCGGAGAGCTGAGGAGAGATGCTGTAACC 1156
----- 1269
1270 ----- 1269
1157 CGGAGGCGGAGT 1216
----- 1269
1270 ----- 1269
1277 GAGATGACACGAGACCTGCTGATGAGAGCGGAGCATGATGATGATGATGATGAT 1336
|||||
1293 GAGATGACACGAGACCTGCTGATGAGAGCGGAGCATGATGATGATGATGATGAT 1352
|||||
1337 CAGCCCCCTACACAGCTGCTTCCGCGAGTTCCGAGCTCTTATGAGGCTTCACG 1396
|||||
1353 CAGCCCCCTACACAGCTGCTTCCGCGAGTTCCGAGCTCTTATGAGGCTTCACG 1412
|||||
1397 TTCACATCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1456
|||||
1413 TTCACATCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1472
|||||
1457 GATGCTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1516
|||||
1473 GATGCTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1532
|||||
1517 CCCAGGCGAACTCAAGAGACAACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1576
|||||
1533 CCCAGGCGAACTCAAGAGACAACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1592
|||||
1577 GGGCTGCGGAG 1636
|||||
1593 GGGCTGCGGAG 1652
|||||
1637 GCGTGAAGCTTAAGGCTGTGAGAGGCTTCACTGAGGAGGAGGAGGAGGAGGAGG 1696
|||||
1653 GCGTGAAGCTTAAGGCTGTGAGAGGCTTCACTGAGGAGGAGGAGGAGGAGGAGG 1712
|||||
1697 TACTACAGTCCCGACAGAGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1756
|||||
1713 TACTACAGTCCCGACAGAGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1772
|||||
1757 TCAGCGCGGTCAAAGCTTCAAGTGTGAGAGGATGAGACACCAAGAGAGAGAGAGAG 1816
|||||
1773 TCAGCGCGGTCAAAGCTTCAAGTGTGAGAGGATGAGACACCAAGAGAGAGAGAGAG 1832
|||||
1817 ACTGTAGTCTGGGCGGAG 1876
|||||
1833 ACTGTAGTCTGGGCGGAG 1892
|||||
1877 ATGAGAGACCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1936
|||||
1893 ATGAGAGACCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1952
|||||
1937 ATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1996
|||||
1953 ATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2012
|||||
1997 CACACTACACTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2056
|||||
2013 CACACTACACTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2072
|||||
2057 CATCTCAACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2116
|||||

Db 2073 CATCTCACTGCTTCTTCTTATGGGAGTTCGCCAGCCAGTCTCACCCTGTGTAC 2132
OY 2117 ACCGACGAGACATGATCCATGATCAGAGCATACAGCTGTCCACACTGAGAGCTGTCT 2176
Db 2133 ACCGACGAGACATGATCCATGATCAGAGCATACAGCTGTCCACACTGAGAGAGTGTCT 2192
OY 2177 ACAACAGCCTGATCAATCAATGTTAGCTTAATAGATATAAATCCAGACTACTTACGCTT 2236
Db 2193 ACAACAGCCTGATCAATCAATGTTAGCTTAATAGATATAAATCCAGACTACTTACGCTT 2252
OY 2237 TAATGCTTTTATTTATATAAATCCGTGAAGTACAGTACATGGAACATTTAACT 2296
Db 2253 TAATGCTTTTATTTATATAAATCCGTGAAGTACAGTACATGGAACATTTAACT 2312
OY 2297 CAGACTGTGATTCAGAGTCCGGAACCCCTTACTTATCTGATCAACAGACACACACC 2356
Db 2313 CAGACTGTGATTCAGAGTCCGGAACCCCTTACTTATCTGATCAACAGACACACACC 2372
OY 2357 TTATGTTACTGCCCCAATATGATTTTAAATATCAATATCTGTTAAAAAATAA 2416
Db 2373 TTATGTTACTGCCCCAATATGATTTTAAATATCAATATCTGTTAAAAAATAA 2432
OY 2417 AAAAAAAAAAAAAA 2429
Db 2433 AAAAAAAAAAAAAA 2445

RESULT 2
BI480798 666 bp mRNA linear EST 28-FEB-2002
LOCUS H2RPE-0430 Human Retinal Pigment Epithelium (2) Homo sapiens cDNA
DEFINITION 5' similar to vitelliform macular dystrophy (Best disease, mRNA
sequence.
VERSION BI480798
KEYWORDS EST.
SOURCE BI480798.1 GI:18998607
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 666)
REFERENCE
AUTHORS Mammals; M., Meers, A.J., Zarepari, S., Farjo, R., Filipova, E.,
Vyan, Y., MacNee, S.P., Hughes, B. and Swaroop, A.
Towards an expression profile of native human retinal pigment
epithelium: Identification of a non-redundant set of more than 1100
genes
JOURNAL
COMMENT unpublished (2001)
CONTACT: Swaroop, A.
Department of Ophthalmology and Visual Sciences
Kellogg Eye Center, University of Michigan
540 KEE, 1000 Wall St., Ann Arbor, MI 48105, USA
Tel: 734 615 2246
Fax: 734 647 0228
Email: swaroop@umich.edu
PCR Primers
FORWARD: M13/PUC-Reverse - ccagtcacagctgttaaacg
BACKWARD: M13/PUC-Forward - agcgataacaattccacacag
Seq primer: M13/PUC-Reverse.

FEATURES
source
1. 666
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone.lib="Human Retinal Pigment Epithelium (2)"
/tissue.type="Native Retinal Pigment Epithelium sheets"
/dev_stage="juvenile"
/note="Organ: Retina; Vector: pSport1"
BASE COUNT 133 a 224 c 171 g 136 t 2 others
ORIGIN

Query Match 26.4%; Score 640.4; DB 13; Length 666;
Best Local Similarity 99.4%; Pore No. 6.7e-85;
Matches 652; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

OY 10 CCACAGCTAGTGCAGACACCTTCTGTGGATCATGAGCCACCCTGGAACCCACCTG 69
Db 12 CCACAGCTAGTGCAGACACCTTCTGTGGATCATGAGCCACCCTGGAACCCACCTG 71
OY 70 ACCCAAGCCACCTGCTGCAGACCCACTGCTGGCCATGACATCACTTACAGAGCAAG 129
Db 72 ACCCAAGCCACCTGCTGCAGACCCACTGCTGGCCATGACATCACTTACAGAGCAAG 131
OY 130 TGGCTATAGCCGCTTAGGCTTCTTCTCCGCTGCTGTGCTGCTGGCGGGGAGCATCT 189
Db 132 TGGCTATAGCCGCTTAGGCTTCTTCTCCGCTGCTGTGCTGCTGGCGGGGAGCATCT 191
OY 190 ACAAGCTGTATATGCGAGTCTTAAATCTTCTGCTGTCTACTACATCACTGCTTAA 249
Db 192 ACAAGCTGTATATGCGAGTCTTAAATCTTCTGCTGTCTACTACATCACTGCTTAA 251
OY 250 TTTTATAGCTGGCCCTCAGGAGAACACAGCTGATGTTGAGAACTGACTCTGATTT 309
Db 252 TTTTATAGCTGGCCCTCAGGAGAACACAGCTGATGTTGAGAACTGACTCTGATTT 311
OY 310 GCGACAGCTACATCAGCTCATCCCATTTCTTCTGCTGTGCTGCTTACTAGTACGCTG 369
Db 312 GCGACAGCTACATCAGCTCATCCCATTTCTTCTGCTGTGCTGCTTACTAGTACGCTG 371
OY 370 TCGTGACCCGCTGGTGGAGACAGTACAGAGACCTGCGTGGCCGACGCTCATGAGCC 429
Db 372 TCGTGACCCGCTGGTGGAGACAGTACAGAGACCTGCGTGGCCGACGCTCATGAGCC 431
OY 430 TGGTGTGGGCTTGTGTCAGAGGACAGGACAGGACGCTGCTGCGGCGACGCTCA 489
Db 432 TGGTGTGGGCTTGTGTCAGAGGACAGGACAGGACGCTGCTGCGGCGACGCTCA 491
OY 490 TCCGCTACGCCAATCCTGGGCAAGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCT 549
Db 492 TCCGCTACGCCAATCCTGGGCAAGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCT 551
OY 550 AGCGCTTCCAGCGCCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
Db 552 AGCGCTTCCAGCGCCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 610
OY 610 ACCAGTGGAGAAATGAGCTTACACACACACACACACACACACACACACACACACAC 666
Db 611 ACCAGTGGAGAAATGAGCTTACACACACACACACACACACACACACACACACACAC 666

RESULT 3
BM663028/c 592 bp mRNA linear EST 27-FEB-2002
LOCUS UI-E-C10-aad-h-10-0-UI-s1 UI-E-C10 Homo sapiens cDNA clone
DEFINITION UI-E-C10-aad-h-10-0-UI 3', mRNA sequence.
ACCESSION BM663028
VERSION BM663028.1 GI:18968017
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 592)
REFERENCE
AUTHORS Bonaldi, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL
COMMENT Genome Res. 6 (9), 791-806 (1996)
MEDLINE
97044477
CONTACT: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Ecksstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msocares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Forward
POLY-A-yes.

FEATURES

source

Location/Qualifiers
1. 592
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-C10-aad-h-10-0-UI"
/clone_1lb="UI-E-C10"
/issue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pRTT3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-C10 is a cDNA library containing the following
tissue(s): RPE and Choroid. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pRTT3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is ACCCTA. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI).
TAG: LIB-UI-E-C10
TAG: LIB-UI-E-C10
TAG-SEQ-ACCCTA"
TAG-SEQ-ACCCTA"

| BASE COUNT | 133 a | 111 c | 142 g | 206 t |
|-----------------------|---|-------|-------|-------|
| ORIGIN | | | | |
| Query Match | 24.0%; Score 582.4; DB 13; Length 592; | | | |
| Best Local Similarity | 99.0%; Pred. No. 2.3e-76; | | | |
| Matches | 586; Conservative 0; Mismatches 6; Indels 0; Gaps 0; | | | |
| QY | 1830 GGGCCAGAAAGTTTGAATTCCTCTCAGAGAGGAGGAGGCTTGTATGAGCAGCCAG 1889 | | | |
| DB | 592 GGGCCAGAAAGTTTGAATTCCTCTCAGAGAGGAGGAGGCTTGTATGAGCAGCCAG 533 | | | |
| QY | 1890 AAGTATCTCAAGTGAAGAGAGAACTGTGAGTTTAACTGACGATATGCGAGATCC 1949 | | | |
| DB | 532 AAGTATCTCAAGTGAAGAGAGAAACCGTGAAGTTTAACTGACGATATGCGAGATCC 473 | | | |
| QY | 1950 CCGAAATATCACTCAAGAAACCTTTGGAACAATCACAACCAACATACATACACTCA 2009 | | | |
| DB | 472 CCGAAATATCACTCAAGAAACCTTTGGAACAATCACAACCAACATACATACACTCA 413 | | | |
| QY | 2010 AAGTATCACTCAAGTGAAGAGAGAACTGTGAGTTTAACTGACGATATGCGAGATCC 2069 | | | |
| DB | 412 AAGTATCACTCAAGTGAAGAGAGAACTGTGAGTTTAACTGACGATATGCGAGATCC 353 | | | |
| QY | 2070 CTTCTCAATGGGAGTGTGCGCCAGCAGAGTCCCTACCTGTGTATACACAGCAGAC 2129 | | | |
| DB | 352 CTTCTCAATGGGAGTGTGCGCCAGCAGAGTCCCTACCTGTGTATACACAGCAGAC 293 | | | |
| QY | 2130 TGATCCAGTCAACAGCCTTACAGTGTCCACACTGAAGAGTGTCTTCAACAAGCTGAA 2189 | | | |
| DB | 292 TGATCCAGTCAACAGCCTTACAGTGTCCACACTGAAGAGTGTCTTCAACAAGCTGAA 233 | | | |
| QY | 2190 TCAATGGTGTAGTTAATAGTAAATAATCCAGACTACTTACGCTTATGCTTTAT 2249 | | | |
| DB | 232 TCAATGGTGTAGTTAATAGTAAATAATCCAGACTACTTACGCTTATGCTTTAT 173 | | | |
| QY | 2250 TCATTAATAAAGTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2309 | | | |
| DB | 172 TCATTAATAAAGTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 113 | | | |
| QY | 2310 CAGAGTGGGAACCTTAGTCTTATGTGAATCCAGAGCAGCAGCTTAGTATATCTGCC 2369 | | | |

| | | |
|----|------|--|
| DB | 112 | CAGAGTGGGAACCTTAGTCTTATGTGAATCCAGAGCAGCAGCTTAGTATATCTGCC 53 |
| QY | 2370 | CAACTATATGAGTTAAT 2421 |
| DB | 52 | CAACTATATGAGTTAAT 1 |

RESULT 4
BE410951
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 773)
NIH-MGC http://mgi.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@rsfemail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: Image.lnl.gov
Plate: LINC338 row: 1 column: 16
High quality sequence stop: 662.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES

source

Location/Qualifiers
1. 773
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3638175"
/clone_1lb="NIH-MGC-21"
/issue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pOT87; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Size-selected
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT

204 a 239 c 176 g 154 t

Query Match 23.8%; Score 577; DB 10; Length 773;
Best Local Similarity 99.0%; Pred. No. 1.3e-75;
Matches 591; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

| | | |
|----|------|---|
| QY | 1456 | GGATGCTACGCTGGCATCATGTCGCTTCTAGCGCTGAGTCCATGATACACATCC 1515 |
| DB | 1 | GGATGCTACGCTGGCATCATGTCGCTTCTAGCGCTGAGTCCATGATACACATCC 60 |
| QY | 1516 | TCCGAGGGGAACTCAAGAGCAACATCACTGTGCGCCAAAGAGGAAATCCCTTTCACGA 1575 |
| DB | 61 | TCCGAGGGGAACTCAAGAGCAACATCACTGTGCGCCAAAGAGGAAATCCCTTTCACGA 120 |
| QY | 1576 | GGGCTGTGCC-AAAAACAAGAGCAAGCAACAAAGAGGAGGAGGAGGAGGAGGAGGAG 1634 |
| DB | 121 | GGGCTGTGCC-AAAAACAAGAGCAAGCAACAAAGAGGAGGAGGAGGAGGAGGAGGAG 180 |
| QY | 1635 | AGGCTGTGAAGCTTAAGGCTGTGAGAGCTTCAAGTGTGCGCCACTGTATCAGAGCCAG 1694 |
| DB | 181 | AGGCTGTGAAGCTTAAGGCTGTGAGAGCTTCAAGTGTGCGCCACTGTATCAGAGCCAG 240 |

QY 1695 GCTACTACAGTGGCCCAAGACGCCCCCTCAGCCCTCAGCTGCTTCCTCCCTAGAAC 1754
 |||||||
 Db 241 GCTACTACAGTGGCCCAAGACGCCCCCTCAGCCCTCAGCTGCTTCCTCCCTAGAAC 300
 |||||||
 QY 1755 CATTGAGCGCGCTCAAGCTTCACAGTGTCTACAGGCGATAGACCAAGACAAAGCTTAA 1814
 |||||||
 Db 301 CATTGAGCGCGCTCAAGCTTCACAGTGTCTACAGGCGATAGACCAAGACAAAGCTTAA 360
 |||||||
 QY 1815 AGACTGTGAGTCTGGGGGCGCCCAAGAAAGTTTGAATGCTCTCAGAGAGCATGGGGCT 1874
 |||||||
 Db 361 AGACTGTGAGTCTGGGGGCGCCCAAGAAAGTTTGAATGCTCTCAGAGAGCATGGGGCT 420
 |||||||
 QY 1875 TGATGAGACACCCAGAAATATCTCAAGTGAAGAGAGAAATCTGTGAGTTTAACCTGACGG 1934
 |||||||
 Db 421 TGATGAGACACCCAGAAATATCTCAAGTGAAGAGAGAAATCTGTGAGTTTAACCTGACGG 480
 |||||||
 QY 1935 ATATGCGAGAGATCCCGGAAATATCACCCTCAAGAAACCTTTGGAGACATCACCACCAACA 1994
 |||||||
 Db 481 ATATGCGAGAGATCCCGGAAATATCACCCTCAAGAAACCTTTGGAGACATCACCACCAACA 540
 |||||||
 QY 1995 TACACACTACACTCAAGATCATGATGATCTTATTTGGGCTTTGGAAACAGAGATG 2051
 |||||||
 Db 541 TACACACTACACTCAAGATCATGATGATCTTATTTGGGCTTTGGAAACAGAGATG 597
 |||||||

RESULT 5 585 bp mRNA linear EST 28-FEB-2002
 LOCUS BM707948
 DEFINITION UI-E-C11-alt-g-11-0-UI-r1 UI-E-C11 Homo sapiens cDNA clone
 UI-E-C11-alt-g-11-0-UI 5', mRNA sequence.
 ACCESSION BM707948
 VERSION BM707948.1 GI:19021206
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 585)
 Ronaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477

JOURNAL MEDLINE
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu

Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.
 Location/Qualifiers

FEATURES

1.585
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-C11-alt-g-11-0-UI"
 /clone_lib="UI-E-C11"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10s (Life Technologies) (T1 phage resistant)"
 /note="Organ: Eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-C11 is a normalized cDNA library containing the
 following tissue(s): RPE and Choroid. The library was
 constructed according to Ronaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dr)18 tail. The sequence tag for this library is ACCGA.
 This library was created for the program, Gene Discovery
 in the Visual System, supported by National Eye Institute
 (NEI).
 BASE COUNT 187 a 145 c 121 g 130 t 2 others
 ORIGIN

Query Match 23.7%; Score 575; DB 14; Length 585;
 Best Local Similarity 98.8%; Pred. No. 2.9e-75;
 Matches 578; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1759 AGCGCGTCAAGCTTCACAGTGTCTCAGAGCATAGACACCAAGACAAAGCTTAAAGAC 1818
 |||||||
 Db 1 AGCGCGTCAAGCTTCACAGTGTCTCAGAGCATAGACACCAAGACAAAGCTTAAAGAC 60
 |||||||
 QY 1819 TGTGAGTCTGGGGGCGCCCAAGAAAGTTTGAATGCTCTCAGAGAGCATGGGGCTTGT 1878
 |||||||
 Db 61 TGTGAGTCTGGGGGCGCCCAAGAAAGTTTGAATGCTCTCAGAGAGCATGGGGCTTGT 120
 |||||||
 QY 1879 GGAGCAGCCAGAAAGATCTCAAGTGAAGAGAGAGAAACGTGGAGTTTAACCTGACGATAT 1938
 |||||||
 Db 121 GGAGCAGCCAGAAAGATCTCAAGTGAAGAGAGAGAGAAACGTGGAGTTTAACCTGACGATAT 180
 |||||||
 QY 1939 GCCAGAGATCCCGGAAATATCACCCTCAAGAAACCTTTGGAGACATCACCACCAACA 1998
 |||||||
 Db 181 GCCAGAGATCCCGGAAATATCACCCTCAAGAAACCTTTGGAGACATCACCACCAACA 240
 |||||||
 QY 1999 CACTTACTCAAGATCATGATGATCTTATTTGGGCTTTGGAAACAGAGATGAAGACA 2058
 |||||||
 Db 241 CACTTACTCAAGATCATGATGATCTTATTTGGGCTTTGGAAACAGAGATGAAGACA 300
 |||||||
 QY 2059 TTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGCTCCTGCTGTGTACAC 2118
 |||||||
 Db 301 TTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGCTCCTGCTGTGTACAC 360
 |||||||
 QY 2119 CAGCAGACACTGATCTCAGTCCAGCCATACAGTGTCTCAGACCTGACAGAAAGCTGTCTAC 2178
 |||||||
 Db 361 CAGCAGACACTGATCTCAGTCCAGCCATACAGTGTCTCAGACCTGACAGAAAGCTGTCTAC 420
 |||||||
 QY 2179 AACAGCCGTGAATCAAAAGTGTAGCTTAATAGATTAATAATCCAGACTACTTACGCTTAA 2238
 |||||||
 Db 421 AACAGCCGTGAATCAAAAGTGTAGCTTAATAGATTAATAATCCAGACTACTTACGCTTAA 480
 |||||||
 QY 2239 ATGCTTTTATCAAAAAAAGCTGAAGCTGAGTGAACCATTTGAACATTTAACTCA 2298
 |||||||
 Db 481 ATGCTTTTATCAAAAAAAGCTGAAGCTGAGTGAACCATTTGAACATTTAACTCA 540
 |||||||
 QY 2299 GACTGTGATTCAGAGTGGGAACCTTATGTTCTATCTGATATCA 2343
 |||||||
 Db 541 GACTGTGATTCAGAGTGGGAACCTTATGTTCTATCTGATATCA 585
 |||||||

RESULT 6 593 bp mRNA linear EST 21-JUL-2000
 LOCUS BE385296
 DEFINITION 60127572F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:361865 5',
 mRNA sequence.
 ACCESSION BE385296
 VERSION BE385296.1 GI:9330661
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 593)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L16M287 row: 0 column: 18
High quality sequence stop: 593.

FEATURES

source

1. 593
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:361865"
/clone_lib="NIH-MGC-20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; site:1: XhoI; site:2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 176 a 177 c 132 g 108 t
ORIGIN

Query Match 23.3%; Score 565.2; DB 10; Length 593;
Best Local Similarity 99.3%; Pred. No. 7.9e-74;
Matches 578; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

1471 CATCATGGGGGCTTCCTAGGCGCTGCAGTCCCATGATACCATCTCTCCAGGCAATC 1530
1 CATATTGGCGGCTTCTAGGCGCTGCAGTCCCATGATACCATCTCTCCAGGCAATC 60
1531 AAGGACCAACTACTGTGGCCCAAGAGGAATCCCTTCCACGAGGCGTGGCC -AAAA 1589
61 AAGGACCAACTACTGTGGCCCAAGAGGAATCCCTTCCACGAGGCGTGGCCCAAAAA 120
1590 ACCCAAGGCGCAACACAGAGCTTAGGGCCAGGAAGACAAAGGCGTGAAGCTTA 1649
121 ACCCAAGGCGCAACACAGAGCTTAGGGCCAGGAAGACAAAGGCGTGAAGCTTA 180
1650 AGGCTGTGAGCGCTTCAAGCTGCGCCACTGTATCAAGGCGGCTACTACAGTGGCC 1709
181 AGGCTGTGAGCGCTTCAAGCTGCGCCACTGTATCAAGGCGGCTACTACAGTGGCC 240
1710 CACAGAGCGCCCTCAGCCCACTGCTCTCCCTGGAACATCAGCGCGTCAA 1769
241 CACAGAGCGCCCTCAGCCCACTGCTCTCCCTGGAACATCAGCGCGTCAA 300
1770 AGCTTCACTGTACACAGGATAGACACCAAGCAAAAGCTTAAGACTGTGATCTG 1829
301 AGCTTCACTGTACACAGGATAGACACCAAGCAAAAGCTTAAGACTGTGATCTG 360
1830 GGGCAAGAAAGTTTGAATGCTCAGAGAGCGATGGGCGCTTATGAGACACCAAG 1889
361 GGGCAAGAAAGTTTGAATGCTCAGAGAGCGATGGGCGCTTATGAGACACCAAG 420
1890 AAGTATCTCAAGTGAAGAGAAACTGTGAGTTTAAGTACAGGATATGACAGATTC 1949
421 AAGTATCTCAAGTGAAGAGAAACTGTGAGTTTAAGTACAGGATATGACAGATTC 480
1950 CCGAAATACCTCCAAAGAACTTTGGAACAATCAGCAACCAACATACACTACACTCA 2009
481 CCGAAATACCTCCAAAGAACTTTGGAACAATCAGCAACCAACATACACTACACTCA 540
2010 AAGATACATGATCTTATTTGGGCTTGGAAAAAGAGATG 2051
541 AAGATACATGATCTTATTTGGGCTTGGAAAAAGAGTCTG 582

RESULT 7

BO879880

LOCUS

DEFINITION

BO879880 963 bp mRNA linear EST 16-ATG-2002
AGENCOURT 8241531 lupskl_dorsal_root_ganglion Homo sapiens CDNA
clone IMAGE:6180559 5', mRNA sequence.

ACCESSION

VERSION

BO879880.1 GI:22271888

KEYWORDS

SOURCE

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L16M13563 row: 1 column: 08
High quality sequence stop: 623.

FEATURES

source

1. 963
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6180559"
/clone_lib="lupskl_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult 36 yr"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6 (Life Technologies); Site:1: NotI; Site:2: SalI; CDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCAGCGCTCG-3' and 5'-GACATGTTCTGATCGGAGCGGCGGCTT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."
BASE COUNT 205 a 289 c 265 g 202 t 2 others
ORIGIN

Query Match 23.2%; Score 564.6; DB 14; Length 963;
Best Local Similarity 77.3%; Pred. No. 7.7e-74;
Matches 795; Conservative 0; Mismatches 31; Indels 203; Gaps 1;

445 TCGAAGCAAGAGACAGACAGCGGCTGTCGCGCCACGCTACCGCCCAAC 504
26 TCGAAGCAAGAGACAGACAGCGGCTGTCGCGCCACGCTACCGCCCAAC 85
505 TGGCAACGTGCTCATCTCTGCGGAGCTGAGACCGGAGCTCAACGCTTCCCAAGC 564
86 TGGCAACGTGCTCATCTCTGCGGAGCTGAGACCGGAGCTCAACGCTTCCCAAGC 145
565 CCGAGCAGCTGTGCAAGAGGCTTATGACTCCGCGAGAAACACAGCTTGGAGAAAC 624
146 CCGAGCAGCTGTGCAAGAGGCTTATGACTCCGCGAGAAACACAGCTTGGAGAAAC 205
625 TGAAGCTACACAAACATCTTGTGGTGGCTTGGTGGTGGTGGTGGTGGTGGTGGT 684
206 TGAAGCTACACAAACATCTTGTGGTGGCTTGGTGGTGGTGGTGGTGGTGGTGGT 265
685 AGGCGTGGCTTGGAGTCAATCCGGAGCCCTATCCGTCGAGAGCGCTGTAAGAGA 744

[illegible]

VERSION B1756228.1 GI:15747806
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 735)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
 Plate: LMNL1487 row: a column: 18
 High quality sequence stop: 577.
 Location/Qualifiers
 1..735

FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5194649"
 /clone_1lb="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb. Insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH-MGC Library."

BASE COUNT 135 a 229 c 214 g 157 t
 ORIGIN

Query Match 21.8%; Score 528.8; DB 13; Length 735;
 Best Local Similarity 92.6%; Pred. No. 1.6e-68;
 Matches 613; Conservative 0; Mismatches 37; Indels 12; Gaps 5;

257 GCTGGCCCTCAGCGAAGAACAGCTGATGTTGAGAACTGACTCTGATTTGGACG 316
 Db 59 GCTGGCCCTCAGCGAAGAACAGCTGATGTTGAGAACTGACTCTGATTTGGACG 118
 Qy 317 CTACATCCAGCTCATCCCATTTCTCTGCTGCTGAGCTGACGCTGCTGCTGAC 376
 Db 119 CTACATCCAGCTCATCCCATTTCTCTGCTGCTGAGCTGACGCTGCTGCTGAC 178
 Qy 377 CCGCTGTGGAACAGTACAGACAGCTGCGTGGCCCGCCGCTCATAGCTGCTGTC 436
 Db 179 CCGCTGTGGAACAGTACAGACAGCTGCGTGGCCCGCCGCTCATAGCTGCTGTC 238
 Qy 437 GGGCTCTCGAAGGCAAGAGGAGGAGGCGGCTGCTGCGGCGACGCTCATCGGCA 496
 Db 239 GGGCTCTCGAAGGCAAGAGGAGGAGGCGGCTGCTGCGGCGACGCTCATCGGCA 298
 Qy 497 CGCCAACTGGGCAACGCTGCTCATCTCTGCGCAGCGCTACAGCAGCTTACAGCGCTT 556
 Db 289 CGCCAACTGGGCAACGCTGCTCATCTCTGCGCAGCGCTACAGCAGCTTACAGCGCTT 358
 Qy 557 CCCAGGCGCCAGCAGCTGCTGCAAGAGGCTTTATGACTCCGCGAAGAACACAGCAGT 616
 Db 359 CCCAGGCGCCAGCAGCTGCTGCAAGAGGCTTTATGACTCCGCGAAGAACACAGCAGT 418
 Qy 617 GGAAGAACTGAGCTTACACACACATGTTCTGGGTGCGCTGGGTGTTGCCAAGCT 676
 Db 419 GGAAGAACTGAGCTTACACACACATGTTCTGGGTGCGCTGGGTGTTGCCAAGCT 478
 Qy 677 GTCAATGAAAGCGTGGGTGAGGTGCAATCCGGACCTTATCTGCTCCAGAGCGCTGCT 736

Db 479 GTCAATGAAAGCGTGGGTGAGGTGCAATCCGGACCTTATCTGCTCCAGAGCTGCT 538
 Qy 737 GAGCAGATGAAACACCTT-GGGTACTAGTGGGACACCTGATATGCTT-ACAGCTGATTT 794
 Db 539 GAGCAGATGAAACACCTTGGGCTTCTGGGTGGACACCTGATATGCTTACAGCTGATTT 598
 Qy 795 AGTATCCACATGGGTATATACAC-----AGTGGTGAATGAGCGGTGATAGCTTCTT 847
 Db 599 AGTATCCACATGGGTATATACAC-----AGTGGTGAATGAGCGGTGATAGCTTCTT 658
 Qy 848 CCTGACTGTCTAGTTGGG--CGGACGTTTGTGA--CCGACCAAGGCTTACCTCGGCCA 904
 Db 659 CCTGACTGTCTAGTTGGGCGGCGGCTTGTGAACCCGCGGAGGCTTCCCTGGGCGC 718
 Qy 905 TG 906
 Db 719 TG 720

RESULT 10
 BM691456 537 bp mRNA linear EST 28-FEB-2002
 LOCUS
 DEFINITION
 UI-E-CII-abe-f-10-0-UI-r1 UI-E-CII Homo sapiens CDNA clone
 ACCESSION
 BM691456
 VERSION
 BM691456.1 GI:19004714
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 COMMENT
 Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seg primer: M13 Reverse.

FEATURES
 source

1..537
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-CII-abe-f-10-0-UI"
 /clone_1lb="UI-E-CII"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)."
 /note="Organ: eye; Vector: pRT73-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-CII is a normalized CDNA library containing the
 following tissue(s): RPE and choroid. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand CDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded CDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pRT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand CDNA contains a library tag
 sequence that is located between the Not I site and the

(df)18 tail. The sequence tag for this library is ACCTA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI)."

BASE COUNT 133 a 172 c 139 g 93 t

Query Match 21.6%; Score 525.8; DB 14; Length 537;
Best Local Similarity 99.6%; Pred. No. 56-68;
Matches 527; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1223 GAGACCACTGATGTCAGAGGAAATTTGAGGTCCTTGGCTGTGATGATG 1282
|||||
8 GAGACCACTGATGTCAGAGGAAATTTGAGGTCCTTGGCTGTGATGATG 67
1283 CACGAGACCTGCTCGGATGAGGCGGACATGTAATGAAATAGCCGAGCAGCC 1342
|||||
68 CACGAGACCTGCTCGGATGAGGCGGACATGTAATGAAATAGCCGAGCAGCC 127
1343 CCGTACACAGCTGCTCGGATGAGGCGGACATGTAATGAAATAGCCGAGCAGCC 1402
128 CCGTACACAGCTGCTCGGATGAGGCGGACATGTAATGAAATAGCCGAGCAGCC 187
1403 ATCAGCTGAAACAAAGAGATGAGTTCAGCCCATCAGAGAGAGAGATGCT 1462
188 ATCAGCTGAAACAAAGAGATGAGTTCAGCCCATCAGAGAGAGAGATGCT 247
1463 CAGCCTGCAATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1522
248 CAGCCTGCAATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
1523 GCAACTCAAGAGACCAACTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1582
308 GCAACTCAAGAGACCAACTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367
1583 CCCAAAACCAAGAGACCAACTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1642
368 CCCAAAACCAAGAGACCAACTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 427
1643 AAGTTAAGCTGTCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1702
428 AAGTTAAGCTGTCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 487
1703 AGTGGCCCAAGAGACCAACTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1751
488 AGTGGCCCAAGAGACCAACTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 536

RESULT 11
BM707649 522 bp mRNA linear EST 28-FEB-2002
LOCUS UI-E-C11-af5-o-15-0-UI-1 UI-E-C11 Homo sapiens cDNA clone
DEFINITION UI-E-C11-af5-o-15-0-UI 5', mRNA sequence.
ACCESSION BM707649
VERSION BM707649.1 GI:19020907
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 522)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msosaresblue.weeg.uiowa.edu

Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers

FEATURES

1..522
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-C11-af5-o-15-0-UI"
/clone_11b="UI-E-C11"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)."
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-C11 is a normalized cDNA library containing the
following tissue(s): RPE and Choroid. The library was
constructed according to Bonaldo, Lennon and Soares,
genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(df)18 tail. The sequence tag for this library is ACCTA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI)."

BASE COUNT 100 a 182 c 130 g 110 t

Query Match 21.5%; Score 522; DB 14; Length 522;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

46 CGGACCCACCTGGAACCCCACTGACCAAGCCACCTGTCAGCCACCTGAGCA 105
1 CGGACCCACCTGGAACCCCACTGACCAAGCCACCTGTCAGCCACCTGAGCA 60
106 TGACCATCATCTTACACAGCCCAAGTGGCTAATGCCGCTTAAGCTCTTCCGCTGC 165
61 TGACCATCATCTTACACAGCCCAAGTGGCTAATGCCGCTTAAGCTCTTCCGCTGC 120
166 TGTGTGCTGAGGCGGAGCATCTACAAAGCTGCTAATGCGAGTTTATCTCTGC 225
121 TGTGTGCTGAGGCGGAGCATCTACAAAGCTGCTAATGCGAGTTTATCTCTGC 180
226 TGTGTGCTGAGGCGGAGCATCTACAAAGCTGCTAATGCGAGTTTATCTCTGC 285
181 TGTGTGCTGAGGCGGAGCATCTACAAAGCTGCTAATGCGAGTTTATCTCTGC 240
286 TGTGTGAGAAACGACCTCTATTGGACAGCTACATCAGCTCATCTCCATCTCTGC 345
241 TGTGTGAGAAACGACCTCTATTGGACAGCTACATCAGCTCATCTCCATCTCTGC 300
346 TGTGTGAGGCTTCTACGTGACGCTGCTGTCAGCCGCTGTCGAGCAAGTACGAACTGC 405
301 TGTGTGAGGCTTCTACGTGACGCTGCTGTCAGCCGCTGTCGAGCAAGTACGAACTGC 360
406 CGTGGCCGACCGCTCATGAGCTGTGTGCGGCTTGTGCAAGGCAAGAGAGCAAG 465
361 CGTGGCCGACCGCTCATGAGCTGTGTGCGGCTTGTGCAAGGCAAGAGAGCAAG 420
466 GCGGCTGTCGAGGCGGAGCATCTACAAAGCTGCTAATGCGAGTTTATCTCTGC 525
421 GCGGCTGTCGAGGCGGAGCATCTACAAAGCTGCTAATGCGAGTTTATCTCTGC 480
526 GCAAGCTGACACCGCAGTCTACAAAGCTTCCCGAGCCGC 567
|||||

Query Match 20.3%; Score 492.4; DB 14; Length 930;
Best Local Similarity 97.4%; Pred. No. 3.2e-63;
Matches 532; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

QY 1189 CATCAACCCCTTTGGAGAGATGATGATTTTGGAGACCAACGATGTCGACAGAA 1248
DB 99 CATCAACCCCTTTGGAGAGATGATGATTTTGGAGACCAACGATGTCGACAGAA 138
QY 1249 TTTGCGAGGTCTCCCTGTTGGCTGTGATGATGACACAGACCTGCTCGATGAGCC 1308
DB 159 TTTGCGAGGTCTCCCTGTTGGCTGTGATGATGACACAGACCTGCTCGATGAGCC 218
QY 1309 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1368
DB 219 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 278
QY 1369 CCGTCGAGCTCTCTTTATGGGCTCACTTCAACATCAGCTGACAAAGAGAGATGA 1428
DB 279 CCGTCGAGCTCTCTTTATGGGCTCACTTCAACATCAGCTGACAAAGAGAGATGA 338
QY 1429 GTTCCAGCCCAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 1488
DB 339 GTTCCAGCCCAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 398
QY 1489 AGGCTGAGCTCTCTTTATGGGCTCACTTCAACATCAGCTGACAAAGAGAGATGA 1548
DB 399 AGGCTGAGCTCTCTTTATGGGCTCACTTCAACATCAGCTGACAAAGAGAGATGA 458
QY 1549 GCCCAAGAGAGAGATGCTTCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 1608
DB 459 GCCCAAGAGAGAGATGCTTCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 518
QY 1609 GAACCTTAGGGGCGATGA 1667
DB 519 GAACCTTAGGGGCGATGA 578
QY 1668 AGCTGGGCGGCACTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 1725
DB 579 AGCTGGGCGGCACTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 638
QY 1726 CCCCC 1731
DB 639 ACCCCC 644

RESULT 14
BM718338 566 bp mRNA linear EST 01-MAR-2002
LOCUS UI-E-E01-a1a-23-0-01.r1 UI-E-E01 Homo sapiens cDNA clone
DEFINITION BM718338
ACCESSION BM718338
VERSION 1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 566)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers

FEATURES
source

1. 566
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-E01-a1a-23-0-01"
/clone_1b="UI-E-E01"
/issue_type="fetal eye"
/dev_stage="fetal"
/lab_host="PH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site: 1: EcoR I; Site: 2: Not I;
UI-E-E01 is a normalized cDNA library containing the
following tissue(s): fetal eye. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(GT)₁₈ tail. The sequence tag for this library is
CCGCTTACCC. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI)."
BASE COUNT 107 a 182 c 145 g 132 t
ORIGIN

Query Match 20.0%; Score 485.2; DB 14; Length 566;
Best Local Similarity 99.4%; Pred. No. 4.5e-62;
Matches 487; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 18 CTATGTCGACAGCTTCTGTGGATATGAGACCACTGGAACCCACCTGACCAAC 77
DB 1 CTATGTCGACAGCTTCTGTGGATATGAGACCACTGGAACCCACCTGACCAAC 60
QY 78 CCACCTGTCGACAGCTTCTGTGGATATGAGACCACTGGAACCCACCTGACCAAC 137
DB 61 CCACCTGTCGACAGCTTCTGTGGATATGAGACCACTGGAACCCACCTGACCAAC 120
QY 138 GCCCGCTTAGGCTCTCTCCGCTGCTGTGCTGCGGGGAGAGATCTACAAGCTG 197
DB 121 GCCCGCTTAGGCTCTCTCCGCTGCTGTGCTGCGGGGAGAGATCTACAAGCTG 180
QY 198 CTATGTCGACAGCTTCTGTGGATATGAGACCACTGGAACCCACCTGACCAAC 257
DB 181 CTATGTCGACAGCTTCTGTGGATATGAGACCACTGGAACCCACCTGACCAAC 240
QY 258 CTGCGCTTCACGATGA 317
DB 241 CTGCGCTTCACGATGA 300
QY 318 TACATTCAGCTTCATCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 377
DB 301 TACATTCAGCTTCATCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 378 CCGTGGTGAACCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 437
DB 361 CCGTGGTGAACCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 420
QY 438 GCGTTCGTCGATGA 497
DB 421 GCGTTCGTCGATGA 480
QY 498 GCCAAGCTGG 507
DB 481 GCCAAGCTGG 490

RESULT 15
BM685122/c
LOCUS
DEFINITION BM685122 503 bp mRNA linear EST 27-FEB-2002
UI-E-EJ1-e}1-1-09-0-UI.s1 UI-E-EJ1 Homo sapiens cDNA clone
UI-E-EJ1-e}1-1-09-0-UI 3', mRNA sequence.
ACCESSION BM685122
VERSION
KEYWORDS BM685122.1 GI:18995018
SOURCE EST.
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 503)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
COMMENT Contact: Soares, MB

Email: msources@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq. Primer: M13 Forward
POLA-yes.

```

FEATURES
source      Location/Qualifiers
1. .503

```

```

BASE COUNT
ORIGIN
96 a 107 c 123 g 177 t

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-EJ1-a)1-1-09-0-U"
/clone_11b="UI-E-EJ1"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foreal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab="Morgan-DH10B (Life Technologies) (T1 phage resistant)"
/note="Host: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site.1: Ecor I; Site.2: Not I;
UI-E-EJ1 is a subtracted cDNA library constructed
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an Ecor I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the syntheses of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dr)18 tail. The
sequence tags for this library are: fetal eyes, AAGATGACGGA
; lens, CCATTAGCGA; eye anterior segment, AATGCGCAT;
optic nerve, CCATTAGTG; retina, CCGCG; Retina foreal and
Macular, GTCC; RPE and Choroid, ACCCTA. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI).
TAG_L1b-UI-E-EJ1
TAG_TISSUE=RPE and Choroid
TAG_SEQ=ACCTA"

```

| | | | | |
|---------------------------|--------|--------------------|-----------|-------------|
| Query Match | 19.6%; | Score 475.4; | DB 14; | length 503; |
| Best Local Similarity | 99.8%; | Pred. No. 1.3e-60; | | |
| Matches 476; Conservative | 0; | Mismatches 1; | Indels 0; | Gaps 0 |

1750 AGAACCATCAGCGCCGTCAAAGCTTCACAGTGTACAGGCATAGACACCAAGACAAG 1809

[illegible]

Search completed: June 9, 2003, 03:30:06
Job time : 3218.69 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 18:10:38 ; Search time 291.723 Seconds

(without alignments)
10693.617 Million cell updates/sec

Title: US-09-622-964-2

Perfect score: 2229

Sequence: 1 caggagcgcaccacacgcta.....aaaaaaaaaaaaaaaaaaaa 2229

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PC7_NEM_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PC7_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|-----------------------|--------------------|
| 1 | 1247.2 | 56.0 | 1263 | 9 US-09-746-783-3 | Sequence 3, Appl1 |
| 2 | 350.4 | 15.7 | 1326 | 9 US-10-071-766-10 | Sequence 10, Appl1 |
| 3 | 215.4 | 9.7 | 1350 | 10 US-09-768-826-16 | Sequence 16, Appl1 |
| 4 | 189.2 | 8.5 | 1292 | 9 US-10-198-846-11070 | Sequence 11070, A |
| 5 | 175 | 7.9 | 1198 | 10 US-09-880-107-2174 | Sequence 2174, A |
| 6 | 116 | 5.2 | 853 | 9 US-10-198-846-11346 | Sequence 11346, A |
| 7 | 104 | 4.7 | 462 | 9 US-10-198-846-2561 | Sequence 2561, A |
| 8 | 102 | 4.6 | 615 | 9 US-10-198-846-7843 | Sequence 7843, A |
| 9 | 98 | 4.4 | 507 | 9 US-09-764-872-195 | Sequence 195, App |
| 10 | 64 | 2.9 | 751 | 9 US-10-198-846-1862 | Sequence 1862, Ap |
| 11 | 49.2 | 2.2 | 1263 | 9 US-09-746-783-3 | Sequence 3, Appl1 |
| 12 | 49.2 | 2.2 | 1326 | 9 US-10-071-766-10 | Sequence 10, Appl1 |
| 13 | 47 | 2.1 | 2076 | 9 US-09-986-480-116 | Sequence 116, Appl |
| 14 | 46.6 | 2.1 | 846 | 10 US-09-349-015-25 | Sequence 25, Appl |
| 15 | 46.6 | 2.1 | 8195 | 9 US-10-156-239-7 | Sequence 7, Appl1 |
| 16 | 46.6 | 2.1 | 8195 | 9 US-10-199-485-7 | Sequence 7, Appl1 |
| 17 | 45.8 | 2.1 | 8195 | 9 US-09-795-693-7 | Sequence 90, Appl |
| 18 | 45.8 | 2.1 | 2338 | 10 US-09-923-302-90 | Sequence 90, Appl |
| 19 | 45.4 | 2.0 | 506 | 9 US-09-918-995-7423 | Sequence 7423, Ap |

| | | | | | |
|----|------|-----|-------|------------------------|-------------------|
| 20 | 45.4 | 2.0 | 667 | 9 US-09-813-153-77 | Sequence 77, Appl |
| 21 | 44.8 | 2.0 | 2341 | 9 US-09-991-053-9 | Sequence 9, Appl1 |
| 22 | 44.8 | 2.0 | 2341 | 9 US-09-957-187-9 | Sequence 9, Appl1 |
| 23 | 44.8 | 2.0 | 2607 | 9 US-09-991-053-11 | Sequence 11, Appl |
| 24 | 44.8 | 2.0 | 2607 | 9 US-09-957-187-11 | Sequence 11, Appl |
| 25 | 44.4 | 2.0 | 4909 | 9 US-10-042-141-25 | Sequence 25, Appl |
| 26 | 44.4 | 2.0 | 4909 | 10 US-09-726-643-25 | Sequence 25, Appl |
| 27 | 44.2 | 2.0 | 827 | 10 US-09-925-301-488 | Sequence 148, App |
| 28 | 44.2 | 2.0 | 912 | 10 US-09-764-853-142 | Sequence 148, App |
| 29 | 44.2 | 2.0 | 913 | 10 US-09-920-300A-1253 | Sequence 348, App |
| 30 | 44 | 2.0 | 400 | 10 US-09-920-300A-1253 | Sequence 1253, Ap |
| 31 | 44 | 2.0 | 400 | 12 US-10-033-528-1253 | Sequence 1253, Ap |
| 32 | 44 | 2.0 | 1342 | 9 US-09-879-389B-1 | Sequence 1, Appl1 |
| 33 | 44 | 2.0 | 15832 | 9 US-10-239-676-106 | Sequence 106, App |
| 34 | 43.8 | 2.0 | 189 | 9 US-10-060-036-2380 | Sequence 2380, Ap |
| 35 | 43.8 | 2.0 | 931 | 9 US-09-991-036-1575 | Sequence 1575, Ap |
| 36 | 43.2 | 1.9 | 442 | 10 US-09-960-352-9688 | Sequence 9688, Ap |
| 37 | 43.2 | 1.9 | 1356 | 9 US-09-951-053-31 | Sequence 31, Appl |
| 38 | 43.2 | 1.9 | 1356 | 9 US-09-957-187-31 | Sequence 31, Appl |
| 39 | 43.2 | 1.9 | 2185 | 9 US-09-982-598-228 | Sequence 228, App |
| 40 | 43.2 | 1.9 | 2185 | 9 US-09-982-293A-228 | Sequence 228, App |
| 41 | 43.2 | 1.9 | 2185 | 9 US-09-982-735-228 | Sequence 228, App |
| 42 | 43.2 | 1.9 | 2185 | 9 US-09-990-444-228 | Sequence 228, App |
| 43 | 43.2 | 1.9 | 2185 | 9 US-09-989-730-228 | Sequence 228, App |
| 44 | 43.2 | 1.9 | 2185 | 9 US-09-989-730-228 | Sequence 228, App |
| 45 | 43.2 | 1.9 | 2185 | 9 US-09-991-181-228 | Sequence 228, App |

ALIGNMENTS

RESULT 1
US-09-746-783-3
Sequence 3, Application US/09746783
Publication No. US20030044935A1

GENERAL INFORMATION:

APPLICANT: JACOBS, Kenneth

McCoy, John M.

Lavalle, Edward R.

Racle, Lisa A.

Treacy, Maurice

Spaulding, Vikki

Agostino, Michael J.

Hoves, Steven H.

Feichtel, Kim

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

ENCODING THEM

NUMBER OF SEQUENCES: 231

CORRESPONDENCE ADDRESSES:

ADDRESS: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

APPLICATION NUMBER: US/09/746,783

FILING DATE: 21-Dec-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Milasincic, Debra J.

REGISTRATION NUMBER: 46,931

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 742-7400

TELEFAX: (617) 742-4214

INFORMATION FOR SEQ ID NO. 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1263 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-746-783-3

Query Match 56.0%; Score 1247.2; DB 9; Length 1263;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 978 GAGCAGCTCATCAACCCCTTGGAGAGATGATGATTTTGGAGACCACTGATGCTC 1037
Db 1 GAGCAGCTCATCAACCCCTTGGAGAGATGATGATTTTGGAGACCACTGATGCTC 60
QY 1038 GACAGGAATTTGCAAGGTGTCCTCTTGGCTGTGATGATGACACAGAGACTGCTCG 1097
Db 61 GACAGGAATTTGCAAGGTGTCCTCTTGGCTGTGATGATGACACAGAGACTGCTCG 120
QY 1098 ATGAGCCGCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1157
Db 121 ATGAGCCGCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
QY 1158 GCCCAGTTCGCTGAGCCCTCTTTATGAGCTCACCCTTCAACATCAGCCTGAACAAAG 1217
Db 181 GCCCAGTTCGCTGAGCCCTCTTTATGAGCTCACCCTTCAACATCAGCCTGAACAAAG 240
QY 1218 GAGATGAGTTCAGCCCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1277
Db 241 GAGATGAGTTCAGCCCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 1278 CGCTTCTTCTAGGCTGAGTCCCATGATCACCATCTCCCAAGGGAACCTCAAGAGCAAA 1337
Db 301 CGCTTCTTCTAGGCTGAGTCCCATGATCACCATCTCCCAAGGGAACCTCAAGAGCAAA 360
QY 1338 CTACTGTGGCCCAAGAGAGATCCCTTCTCCAGAGGCGCTGCCCAAAACCAAGGCA 1397
Db 361 CTACTGTGGCCCAAGAGAGATCCCTTCTCCAGAGGCGCTGCCCAAAACCAAGGCA 420
QY 1398 GCCAAACAGAACGTTAGGGGCGCAAGAACACAAAGCCGGAAGCTTAAGGCTGTGAC 1457
Db 421 GCCAAACAGAACGTTAGGGGCGCAAGAACACAAAGCCGGAAGCTTAAGGCTGTGAC 480
QY 1458 GCCTTCAAGTCTGGCCCATGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1517
Db 481 GCCTTCAAGTCTGGCCCATGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 1518 CTGAGGCCCACTGCTGCTTCTTCCCTTGAACCATCAGCGCGCTCAAAAGCTTCAAG 1577
Db 541 CTGAGGCCCACTGCTGCTTCTTCCCTTGAACCATCAGCGCGCTCAAAAGCTTCAAG 600
QY 1578 GTACACAGCATAGACACCAAAAGACCTTAAGAGCTGTGATCTGGGGCCCAAGAA 1637
Db 601 GTACACAGCATAGACACCAAAAGACCTTAAGAGCTGTGATCTGGGGCCCAAGAA 660
QY 1638 AGTTTGAATTTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1697
Db 661 AGTTTGAATTTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 1698 GTGAGAGAGAAACTGTGAGTTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1757
Db 721 GTGAGAGAGAAACTGTGAGTTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 1758 CTCAAAGAACCTTTGGAAACATCACCACCAACATCAGACTCACTCAAGATCAGAT 1817
Db 781 CTCAAAGAACCTTTGGAAACATCACCACCAACATCAGACTCACTCAAGATCAGAT 840
QY 1818 GATCCTTATTTGGGCTTGGAAACAGGAGATGAAGCATTTCTTAACCTGCTTCTTAAG 1877
Db 841 GATCCTTATTTGGGCTTGGAAACAGGAGATGAAGCATTTCTTAACCTGCTTCTTAAG 900
QY 1878 GGATGCTTCGCGAGGAGGCTCTCAGCTGTGTACACAGAGAGAGAGAGAGAGAGAG 1937
Db 1878 GGATGCTTCGCGAGGAGGCTCTCAGCTGTGTACACAGAGAGAGAGAGAGAGAGAG
```

```
Db 901 GGATGCTTCGCGAGGAGGCTCTCAGCTGTGTGTACACAGAGAGAGAGAGAGAGAG 960
QY 1938 CAGCCATACAGCTGTCCACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1997
Db 961 CAGCCATACAGCTGTCCACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1998 GCTTAAATAGATTAATAATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2057
Db 1021 GCTTAAATAGATTAATAATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 2058 GTGAAAGCTAGACTGAACCAATTTGAAACATTTAATCAGACTGTGATTCAGAGT 2117
Db 1081 GTGAAAGCTAGACTGAACCAATTTGAAACATTTAATCAGACTGTGATTCAGAGT 1140
QY 2118 ACCCTTGTCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 2177
Db 1141 ACCCTTGTCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 1200
QY 2178 GTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2229
Db 1201 GTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1252
```

RESULT 2

US-10-071-766-10/c
Sequence 10, Application US/10071766
Publication No. US20020192678A1
GENERAL INFORMATION:
APPLICANT: Huel-Mei Chen
TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
FILE REFERENCE: PA-0043 US
CURRENT APPLICATION NUMBER: US/10/071,766
CURRENT FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 144
SOFTWARE: PERL Program
SEQ ID NO 10
LENGTH: 1326
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. US20020192678A1 1138151.2
US-10-071-766-10

Query Match 15.7%; Score 350.4; DB 9; Length 1326;
Best Local Similarity 99.7%; Pred. No. 2,8e-93;
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1850 AGCAGCTTCTTACCGCTTCTTAAATGGGAGATGCTTGGCAGCAGAGTCTTCACTGTGT 1909
Db 1326 AGCAGCTTCTTACCGCTTCTTAAATGGGAGATGCTTGGCAGCAGAGTCTTCACTGTGT 1267
QY 1910 GTACACAGCAGAGACAGTATCAGTACAGCAGCAGTATCAGAGTGTCCACACTGAAGAGTG 1969
Db 1266 GTACACAGCAGAGACAGTATCAGTACAGCAGCAGTATCAGAGTGTCCACACTGAAGAGTG 1207
QY 1970 TCTTCAACAGAGCTGATCAATTAATGTTAGCTTAATTAATTAATTAATTAATTAATTA 2029
Db 1206 TCTTCAACAGAGCTGATCAATTAATGTTAGCTTAATTAATTAATTAATTAATTAATTA 1147
QY 2030 CCTTAAATGCTTATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2089
Db 1146 CCTTAAATGCTTATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1087
QY 2090 AACTCAGACTGTGATTCAGAGTGGGAAACCTTATGTTCTATTCAGATCAAGAGAGCA 2149
Db 1086 AACTCAGACTGTGATTCAGAGTGGGAAACCTTATGTTCTATTCAGATCAAGAGAGCA 1027
QY 2150 CACCTTATATACAGCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2201
Db 1026 CACCTTATATACAGCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 975
```

```

1  LOCATION: (1345)
2  OTHER INFORMATION: n equals a,t,g, or c
3  NAME/KEY: SITE
4  LOCATION: (1348)
5  OTHER INFORMATION: n equals a,t,g, or c
6  US-09-768-826-16
7
8  Query Match          9 7%; Score 215.4; DB 10; Length 1350;
9  Best Local Similarity 61.9%; Prid. No. 3,8e-53;
10 Matches 420; Conservative 0; Mismatches 211; Indels 48; Gaps 3
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 10:29:42 ; Search time 100.098 Seconds
(without alignments)
7441.920 Million cell updates/sec

Title: US-09-622-964-4
2429
Sequence: 1 caggagatccaccagccta.....aaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/prodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/prodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/prodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/prodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/prodata/1/lna/ECTUS.COMB.seq:*
6: /cgn2_6/prodata/1/lna/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------------|-------------------|
| C 1 | 84.6 | 3.5 | 5590 | 4 US-09-050-159-129 | Sequence 129, App |
| C 2 | 84.6 | 3.5 | 9365 | 4 US-09-608-285A-8 | Sequence 8, Appl |
| C 3 | 84.6 | 3.5 | 9365 | 4 US-09-350-835B-8 | Sequence 8, Appl |
| C 4 | 84.6 | 3.5 | 9365 | 4 US-09-370-265-8 | Sequence 8, Appl |
| C 5 | 84.6 | 3.5 | 14747 | 4 US-09-608-285A-42 | Sequence 42, Appl |
| C 6 | 84.6 | 3.5 | 15977 | 4 US-09-608-285A-59 | Sequence 59, Appl |
| C 7 | 84.6 | 3.5 | 50000 | 4 US-09-146-053-3 | Sequence 3, Appl |
| C 8 | 84.2 | 3.5 | 11531 | 1 US-08-068-945A-1 | Sequence 1, Appl |
| C 9 | 84.2 | 3.5 | 11531 | 1 US-08-442-806-1 | Sequence 1, Appl |
| C 10 | 83.2 | 3.4 | 36741 | 4 US-09-301-665-3 | Sequence 3, Appl |
| C 11 | 83 | 3.4 | 1701 | 4 US-09-078-294-9 | Sequence 9, Appl |
| C 12 | 83 | 3.4 | 2115 | 4 US-08-395-800A-7 | Sequence 7, Appl |
| C 13 | 83 | 3.4 | 6769 | 1 US-08-480-784-20 | Sequence 20, Appl |
| C 14 | 83 | 3.4 | 6769 | 1 US-08-483-553-20 | Sequence 20, Appl |
| C 15 | 83 | 3.4 | 6769 | 1 US-08-487-002-20 | Sequence 20, Appl |
| C 16 | 83 | 3.4 | 6769 | 1 US-08-483-554B-20 | Sequence 20, Appl |
| C 17 | 83 | 3.4 | 6769 | 1 US-08-488-011B-20 | Sequence 20, Appl |
| C 18 | 83 | 3.4 | 6769 | 4 US-08-850-727-20 | Sequence 20, Appl |
| C 19 | 83 | 3.4 | 6769 | 5 PCT-US95-10202-20 | Sequence 20, Appl |
| C 20 | 83 | 3.4 | 6769 | 5 PCT-US95-10203-20 | Sequence 20, Appl |
| C 21 | 83 | 3.4 | 6769 | 5 PCT-US95-10220-20 | Sequence 20, Appl |
| C 22 | 83 | 3.4 | 14636 | 4 US-09-173-914-6 | Sequence 6, Appl |
| C 23 | 83 | 3.4 | 20674 | 4 US-09-641-638-651 | Sequence 651, App |
| C 24 | 83 | 3.4 | 70000 | 4 US-09-851-896-3 | Sequence 3, Appl |
| C 25 | 83 | 3.4 | 112132 | 4 US-09-741-150-3 | Sequence 3, Appl |
| C 26 | 82.6 | 3.4 | 282 | 1 US-08-133-629-8 | Sequence 8, Appl |
| C 27 | 82.6 | 3.4 | 1613 | 2 US-08-812-204-1 | Sequence 1, Appl |

| | | | | | |
|------|------|-----|--------|---------------------|-------------------|
| C 28 | 82.6 | 3.4 | 11725 | 2 US-08-756-506-1 | Sequence 1, Appl |
| C 29 | 82.6 | 3.4 | 38564 | 4 US-09-734-673-3 | Sequence 3, Appl |
| C 30 | 82 | 3.4 | 59065 | 4 US-09-813-817-3 | Sequence 3, Appl |
| C 31 | 82 | 3.4 | 59065 | 4 US-09-978-197-3 | Sequence 3, Appl |
| C 32 | 82 | 3.4 | 99500 | 4 US-09-798-096-10 | Sequence 10, Appl |
| C 33 | 82 | 3.4 | 111282 | 4 US-09-754-250-3 | Sequence 3, Appl |
| C 34 | 81.4 | 3.4 | 4421 | 2 US-08-257-963B-9 | Sequence 9, Appl |
| C 35 | 81.4 | 3.4 | 4421 | 4 US-08-367-841A-6 | Sequence 6, Appl |
| C 36 | 81.4 | 3.4 | 4421 | 4 US-08-520-373D-6 | Sequence 6, Appl |
| C 37 | 81.4 | 3.4 | 4421 | 5 PCT-US95-07201-9 | Sequence 1, Appl |
| C 38 | 81.4 | 3.4 | 5581 | 4 US-08-973-544-1 | Sequence 1, Appl |
| C 39 | 81.4 | 3.4 | 5789 | 4 US-09-242-948-3 | Sequence 3, Appl |
| C 40 | 81.4 | 3.4 | 7705 | 2 US-08-687-080-115 | Sequence 115, App |
| C 41 | 81.4 | 3.4 | 12394 | 4 US-09-488-856A-10 | Sequence 10, Appl |
| C 42 | 81.4 | 3.4 | 20598 | 4 US-09-593-995-10 | Sequence 10, Appl |
| C 43 | 81.4 | 3.4 | 36159 | 4 US-09-743-588-3 | Sequence 3, Appl |
| C 44 | 81.4 | 3.4 | 45716 | 4 US-08-965-048-5 | Sequence 5, Appl |
| C 45 | 81.4 | 3.4 | 45989 | 4 US-08-965-048-6 | Sequence 6, Appl |

ALIGNMENTS

```

RESULT 1
US-09-050-159-129/c
Sequence 129, Application US/09050159A
Patent No. 6197505
GENERAL INFORMATION:
APPLICANT: No. 6197505berg, Ielf T
APPLICANT: Andersson, Maria K
APPLICANT: Ilnström, Per H
TITLE OF INVENTION: METHODS FOR ASSESSING CARDIOVASCULAR STATUS AND
FILE REFERENCE: 1248/10042
CURRENT APPLICATION NUMBER: US/09/050,159A
CURRENT FILING DATE: 1998-03-27
EARLIER APPLICATION NUMBER: 60/042,930
EARLIER FILING DATE: 1987-04-03
NUMBER OF SEQ ID NOS: 133
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 129
LENGTH: 5590
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Angiotensin I converting enzyme, 5' region
US-09-050-159-129
Query Match 3.5%; Score 84.6; DB 4; Length 5590;
Best Local Similarity 90.9%; Pred. No. 5.9e-11;
Matches 90; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1079 AGAAGCAGCCAGCGGTGTGGCGACACCTGTATCCAGCTACTCGGAGGCTGAGCC 1138
D 1926 AAAAAATTAGCAGCGGTGTGGCGATCTCTTAATCCAGCTACTCGGAGGCTGAGCC 1867
QY 1139 AGAGAATTCCTTGAACCCGGAGGCGAGGAGCTGTGTG 1177
D 1866 AGAGAATTCCTTGAACCCGGAGGCGAGAGATTGTGTG 1828
RESULT 2
US-09-608-285A-8/c
Sequence 8, Application US/09608285A
Patent No. 6335013
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
APPLICANT: Young, George
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608,285A

```

```

CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/583,231
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557,800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481,238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244,444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122,449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 9365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (3409)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
NAME/KEY: misc.feature
LOCATION: (9214)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
NAME/KEY: misc.feature
LOCATION: (9303)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
NAME/KEY: misc.feature
LOCATION: (9311)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-09-608-285A-8

Query Match
Best Local Similarity 3.5%; Score 84.6; DB 4; Length 9365;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGACGACGCGGTGGGCGACACCTGTATCCAGCTACTCGGAGGCTGAGGC 1138
DB 7598 AAAAATTAGCCGCGGTGGGCGACCTGTATCCAGCTACTCTGGAGGCTGAGGC 7539
QY 1139 AGAGATCGCTTGAACCCGCGGAGCGGAGGCTGTGTG 1177
DB 7538 AGRAGATCGCTTGAACCCGAGAGGAGGTGCACTG 7500

RESULT 3
US-09-350-836B-8/C
Sequence 8, Application US/09350836B
Patent No. 6387645
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
FILE REFERENCE: 28110/35761
CURRENT APPLICATION NUMBER: US/09/350,836B
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-16
PRIOR APPLICATION NUMBER: 09/122,449
PRIOR FILING DATE: 1998-07-24
```

```

PRIOR APPLICATION NUMBER: 09/244,444
PRIOR FILING DATE: 1998-02-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 9365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: exon
LOCATION: (1) .. (288)
NAME/KEY: exon
LOCATION: (1281) .. (1580)
NAME/KEY: exon
LOCATION: (1820) .. (1855)
NAME/KEY: exon
LOCATION: (2467) .. (2555)
NAME/KEY: exon
LOCATION: (2863) .. (2942)
NAME/KEY: exon
LOCATION: (3889) .. (3950)
NAME/KEY: exon
LOCATION: (4894) .. (4995)
NAME/KEY: exon
LOCATION: (5847) .. (5987)
NAME/KEY: exon
LOCATION: (6966) .. (7138)
NAME/KEY: exon
LOCATION: (8556) .. (9365)
NAME/KEY: misc.feature
LOCATION: (3409)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
NAME/KEY: misc.feature
LOCATION: (9214)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
NAME/KEY: misc.feature
LOCATION: (9303)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
NAME/KEY: misc.feature
LOCATION: (9311)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
US-09-350-836B-8

Query Match
Best Local Similarity 3.5%; Score 84.6; DB 4; Length 9365;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGACGACGCGGTGGGCGACACCTGTATCCAGCTACTCGGAGGCTGAGGC 1138
DB 7598 AAAAATTAGCCGCGGTGGGCGACCTGTATCCAGCTACTCTGGAGGCTGAGGC 7539
QY 1139 AGAGATCGCTTGAACCCGCGGAGCGGAGGCTGTGTG 1177
DB 7538 AGRAGATCGCTTGAACCCGAGAGGAGGTGCACTG 7500

RESULT 4
US-09-370-265-8/C
Sequence 8, Application US/09370265
Patent No. 6447771
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
FILE REFERENCE: 28111/35908
CURRENT APPLICATION NUMBER: US/09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
```

EARLIER FILING DATE: 1999-03-19
EARLIER APPLICATION NUMBER: 09/244,444
EARLIER FILING DATE: 1999-02-04
EARLIER APPLICATION NUMBER: 09/122,449
EARLIER FILING DATE: 1998-07-24
EARLIER APPLICATION NUMBER: 09/118,205
EARLIER FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8

LENGTH: 9365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (72)..(287)
FEATURE:
NAME/KEY: CDS
LOCATION: (1280)..(1579)
FEATURE:
NAME/KEY: CDS
LOCATION: (1819)..(1854)
FEATURE:
NAME/KEY: CDS
LOCATION: (2466)..(2555)
FEATURE:
NAME/KEY: CDS
LOCATION: (2863)..(2940)
FEATURE:
NAME/KEY: CDS
LOCATION: (3887)..(3952)
FEATURE:
NAME/KEY: CDS
LOCATION: (4896)..(4994)
FEATURE:
NAME/KEY: CDS
LOCATION: (5846)..(5986)
FEATURE:
NAME/KEY: CDS
LOCATION: (6965)..(7138)
FEATURE:
NAME/KEY: CDS
LOCATION: (8556)..(8639)
FEATURE:
NAME/KEY: misc feature
LOCATION: (3409)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
FEATURE:
NAME/KEY: misc feature
LOCATION: (9214)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
FEATURE:
NAME/KEY: misc feature
LOCATION: (9303)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
FEATURE:
NAME/KEY: misc feature
LOCATION: (9311)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
US-09-370-265-8

Query Match 3.5%; Score 84.6; DB 4; Length 9365;
Best Local Similarity 84.8%; Pred. No. 7.3e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGACGACGCGGTGTCGCGCACACCTGTATCCAGCTACTCGGAGGCTGAGGC 1138
DB 7598 AAAAATTACCGCGCGTGTGCGCGCGCTGTATCCAGCTACTCGGAGGCTGAGGC 7539
QY 1139 AGAGAAATCGCTTGAACCCGCGGAGCGGAGGCTGTGTG 1177
DB 7538 AGAGAAATCGCTTGAACCCAGGAGGAGGAGGCTGTG 7500

RESULT 5
US-09-608-285A-42/c
Sequence 42, Application US/09608285A
Patent No. 6335013

GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
APPLICANT: Yeung, George
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608,285A
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/583,231
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557,800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481,238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244,444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122,449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 42
LENGTH: 14747
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (13641)
OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine
US-09-608-285A-42

Query Match 3.5%; Score 84.6; DB 4; Length 14747;
Best Local Similarity 84.8%; Pred. No. 8.9e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGACGACGCGGTGTCGCGCACACCTGTATCCAGCTACTCGGAGGCTGAGGC 1138
DB 10958 AAAAATTACCGCGCGTGTGCGCGCGCTGTATCCAGCTACTCGGAGGCTGAGGC 10899
QY 1139 AGAGAAATCGCTTGAACCCGCGGAGCGGAGGCTGTGTG 1177
DB 10898 AGAGAAATCGCTTGAACCCAGGAGGAGGAGGCTGTG 10860

RESULT 6
US-09-608-285A-59/c
Sequence 59, Application US/09608285A
Patent No. 6335013
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
APPLICANT: Yeung, George
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608,285A
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/583,231

PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557,800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481,238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244,444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122,449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-16
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 59
LENGTH: 15977
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: CD39-L4/L66 Gene Sequence

NAME/KEY: CDS
LOCATION: (245)..(461)
NAME/KEY: CDS
LOCATION: (1454)..(1533)
NAME/KEY: CDS
LOCATION: (2734)..(2877)
NAME/KEY: CDS
LOCATION: (4364)..(4439)
NAME/KEY: CDS
LOCATION: (4679)..(4714)
NAME/KEY: CDS
LOCATION: (5326)..(5414)
NAME/KEY: CDS
LOCATION: (5723)..(5802)
NAME/KEY: CDS
LOCATION: (6751)..(6812)
NAME/KEY: CDS
LOCATION: (7758)..(7859)
NAME/KEY: CDS
LOCATION: (8712)..(8852)
NAME/KEY: CDS
LOCATION: (9831)..(9887)
NAME/KEY: CDS
LOCATION: (11613)..(11728)
NAME/KEY: CDS
LOCATION: (13146)..(13691)
NAME/KEY: CDS
LOCATION: (15702)..(15839)
NAME/KEY: misc_feature
LOCATION: (14871)
OTHER INFORMATION: n - a or c or g or t
US-09-608-285A-59

Query Match 3.5%; Score 84.6; DB 4; Length 15977;
Best Local Similarity 84.8%; Fred. No. 9.2e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;
Db 1079 AGAAGCAGCAGCGTGTGGCGCACCTGTATCCAGCTCTGGAGGCGGAGGC 1138
12188 AAAAATAGCGCGGCGGTGTGGCGCACCTGTATCCAGCTCTGGAGGCGGAGGC 12129
QY 1139 AGGAGATCGCTTGAACCGGAGCGGAGGCTGTGTGTG 1177
12128 AGGAGATCGCTTGAACCGGAGCGGAGGCTGTGTGTG 12090

RESULT 7
US-09-146-053-3/c
Sequence 3, Application US/09146053A
Patent No. 639349
GENERAL INFORMATION:
APPLICANT: Ryan, James W.
APPLICANT: Sprinkle, Terry Joe Curtis
APPLICANT: Venema, Richard C.
TITLE OF INVENTION: Human Aminopeptidase P Gene
FILE REFERENCE: MGI03
CURRENT APPLICATION NUMBER: US/09/146,053A
CURRENT FILING DATE: 1998-09-02
EARLIER APPLICATION NUMBER: 60/057,854
EARLIER FILING DATE: 1997-09-02
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 3
LENGTH: 50000
TYPE: DNA
ORGANISM: Homo sapiens
US-09-146-053-3

Query Match 3.5%; Score 84.6; DB 4; Length 50000;
Best Local Similarity 90.9%; Fred. No. 1.5e-10;
Matches 90; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Db 1079 AGAAGCAGCAGCGTGTGGCGCACCTGTATCCAGCTCTGGAGGCGGAGGC 1138
8689 AAAAATAGCGCGGCGGTGTGGCGCACCTGTATCCAGCTCTGGAGGCGGAGGC 8630
QY 1139 AGGAGATCGCTTGAACCGGAGCGGAGGCTGTGTGTG 1177
Db 8629 AGGAGATCGCTTGAACCGGAGCGGAGGCTGTGTGTG 8591

RESULT 8
US-08-068-945A-1
Sequence 1, Application US/08068945A
Patent No. 5616483
GENERAL INFORMATION:
APPLICANT: Bjursell, Gunnar
APPLICANT: Carlsson, Peter
APPLICANT: Eneback, Sven
APPLICANT: Hansson, Lennart
APPLICANT: Lidberg, Ulf
APPLICANT: Nilsson, Jeanette
APPLICANT: Tornell, Jan
TITLE OF INVENTION: New DNA Sequences
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/068,945A
FILING DATE: 27-MAY-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9201809-2
FILING DATE: 11-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9201826-6
FILING DATE: 12-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9202088-2

FILING DATE: 03-JUL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9300902-5
 FILING DATE: 19-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Sterner, Richard J.
 REGISTRATION NUMBER: 35,372
 REFERENCE/DOCKET NUMBER: 1103326-052
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)819-8783
 TELEFAX: (212)354-8113
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 11531 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 TISSUE TYPE: Mammary gland
 FEATURE:
 NAME/KEY: CDS
 LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
 LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
 LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11394)
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707
 LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
 LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
 OTHER INFORMATION: /EC_number= 3.1.1.1
 FEATURE:
 NAME/KEY: 5'UTR
 LOCATION: 1..1640
 FEATURE:
 NAME/KEY: TATA_signal
 LOCATION: 1611..1617
 FEATURE:
 NAME/KEY: exon
 LOCATION: 1641..1727
 FEATURE:
 NAME/KEY: exon
 LOCATION: 4071..4221
 FEATURE:
 NAME/KEY: exon
 LOCATION: 4307..4429
 FEATURE:
 NAME/KEY: exon
 LOCATION: 4707..4904
 FEATURE:
 NAME/KEY: exon
 LOCATION: 6193..6323
 FEATURE:
 NAME/KEY: exon
 LOCATION: 6501..6608
 FEATURE:
 NAME/KEY: exon
 LOCATION: 6751..6868
 FEATURE:
 NAME/KEY: exon
 LOCATION: 8335..8521
 FEATURE:
 NAME/KEY: exon
 LOCATION: 8719..8922
 FEATURE:
 NAME/KEY: exon
 LOCATION: 10124..10321
 FEATURE:
 NAME/KEY: exon
 LOCATION: 10650..11490
 FEATURE:

NAME/KEY: 3'UTR
 LOCATION: 11491..11531
 US-08-068-945A-1
 Query Match
 Best Local Similarity 91.8%; Score 84.2; DB 1; Length 11531;
 Matches 89; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1081 AAAGCAGCAGCGCTGTGCGCCACACCTGTATCCCACTACTCGGAGGCTGAGCAG 1140
 DB 5361 AATTAAGCAGCGCTGTGCGCCCTTGCCCTGTATCCCACTACTCGGAGGCTGAGCAG 5420
 QY 1141 GAGATCGCTTGAACCCGGAGCGGAGGCTGTGCTG 1177
 DB 5421 GAGATCGCTTGAACCTGACGAGCGGAGGCTGCGGTG 5457
 RESULT 9
 US-08-442-806-1
 Sequence 1, Application US/08442806
 Patent No. 5716817
 GENERAL INFORMATION:
 APPLICANT: Bjursell, Gunnar
 APPLICANT: Carlsson, Peter
 APPLICANT: Enerback, Sven
 APPLICANT: Hansson, Lennart
 APPLICANT: Lidberg, Ulf
 APPLICANT: Nilsson, Jeanette
 APPLICANT: Tornell, Jan
 TITLE OF INVENTION: Genomic DNA Sequences
 TITLE OF INVENTION: Encoding Human BSSL/CEL
 NUMBER OF SEQUENCES: 58
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: White & Case
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States
 ZIP: 10036-2787
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/442,806
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/068,945
 FILING DATE: 27-MAY-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9201809-2
 FILING DATE: 11-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9201826-6
 FILING DATE: 12-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9202088-2
 FILING DATE: 03-JUL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9300902-5
 FILING DATE: 19-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Sterner, Richard J.
 REGISTRATION NUMBER: 35,372
 REFERENCE/DOCKET NUMBER: 1103326-052
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)819-8783
 TELEFAX: (212)354-8113
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:

LENGTH: 11531 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Mammary gland
FEATURE:
NAME/KEY: CDS
LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
FEATURE:
NAME/KEY: mat.peptide
LOCATION: join(1727..1727, 4071..4221, 4307..4429, 4707
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
OTHER INFORMATION: /EC_number=3.1.1.1
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..1640
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 1611..1617
FEATURE:
NAME/KEY: exon
LOCATION: 1641..1727
FEATURE:
NAME/KEY: exon
LOCATION: 4071..4221
FEATURE:
NAME/KEY: exon
LOCATION: 4307..4429
FEATURE:
NAME/KEY: exon
LOCATION: 4707..4904
FEATURE:
NAME/KEY: exon
LOCATION: 6193..6323
FEATURE:
NAME/KEY: exon
LOCATION: 6501..6608
FEATURE:
NAME/KEY: exon
LOCATION: 6751..6868
FEATURE:
NAME/KEY: exon
LOCATION: 8335..8521
FEATURE:
NAME/KEY: exon
LOCATION: 8719..8922
FEATURE:
NAME/KEY: exon
LOCATION: 10124..10321
FEATURE:
NAME/KEY: exon
LOCATION: 10650..11490
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 11491..11531
US-08-442-806-1

Query Match 3.5%; Score 84.2; DB 1; Length 11531;
Best Local Similarity 91.8%; Pred. No. 1e-10;
Matches 89; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1081 AAAGCAGCCAGCGCTGGTGGCGCACACCTGTAAATCCAGCTACTCGGAGGCTGAGGCAG 1140
DB 5361 AAATTAGCCAGCGCTGGTGGCGCTGTGCTGTAAATCCAGCTACTCGGAGGCTGAGGCAG 5420
QY 1141 GAGAAATCGCTTGAACCCGGAGGCGGAGGCTTGTGCTG 1177

DB 5421 GAGAAATCGCTTGAACCTCAGGAGGCGGAGGCTTGCCTG 5457
RESULT 10
US-09-301-665-3/c
Sequence 3, Application US/09301665
Patent No. 6207876
GENERAL INFORMATION:
APPLICANT: KELLEMS, RODNEY E.
APPLICANT: DATTA, SURJIT K.
APPLICANT: BLACKBURN, MICHAEL R.
TITLE OF INVENTION: ADENOSINE DEAMINASE DEFICIENT TRANSGENIC MICE AND
FILE REFERENCE: METHODS FOR THE USE THEREOF
CURRENT APPLICATION NUMBER: US/09/301,665
CURRENT FILING DATE: 1999-04-28
EARLIER APPLICATION NUMBER: 60/083,408
EARLIER FILING DATE: 1998-04-29
EARLIER APPLICATION NUMBER: 60/083,370
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 36741
TYPE: DNA
ORGANISM: Homo sapiens
US-09-301-665-3

Query Match 3.4%; Score 83.2; DB 4; Length 36741;
Best Local Similarity 80.8%; Pred. No. 2.9e-10;
Matches 97; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1081 AAAGCAGCCAGCGCTGGTGGCGCACACCTGTAAATCCAGCTACTCGGAGGCTGAGGCAG 1140
DB 1531 AAATTAGCCAGCGCTGGTGGCTGTATGACTGTAAATCCAGCTACTCGGAGGCTGAGGCAG 1472
QY 1141 GAGAAATCGCTTGAACCCGGAGGCGGAGGCTTGTGCTGAGGAGGCTTGTGCTG 1200
DB 1471 GAGAAATCGCTTGAACCTCAGGAGGCTTGTGCTGAGGAGGCTTGTGCTGAGGAGGCTTGTGCTG 1412

RESULT 11
US-09-078-294-9/c
Sequence 9, Application US/09078294
Patent No. 6265211
GENERAL INFORMATION:
APPLICANT: Choo, Kong-Hong Andy
APPLICANT: Du Sart, Desiree
APPLICANT: Cancilla, Michael R.
TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
FILE REFERENCE: Davies Col
CURRENT APPLICATION NUMBER: US/09/078,294
CURRENT FILING DATE: 1998-05-13
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 1701
TYPE: DNA
ORGANISM: BAC-F2 contig 5
US-09-078-294-9

Query Match 3.4%; Score 83; DB 4; Length 1701;
Best Local Similarity 89.9%; Pred. No. 8.6e-11;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1079 AGAAGCAGCCAGCGCTGGTGGCGCACACCTGTAAATCCAGCTACTCGGAGGCTGAGGC 1138
DB 1274 AAATTAGCCAGCGCTGGTGGCTGTGCGCATGCTGTAAATCCAGCTACTCGGAGGCTGAGGC 1215
QY 1139 AGAAGCAGCCAGCGCTGGTGGCGCACACCTGTAAATCCAGCTACTCGGAGGCTGAGGC 1177
DB 1214 AGAAGCAGCCAGCGCTGGTGGCGCACACCTGTAAATCCAGCTACTCGGAGGCTGAGGC 1176

RESULT 12
US-08-395-800A-7
Sequence 7, Application US/08395800A
Patent No. 5807732
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B
APPLICANT: LENNON, GREGORY
APPLICANT: ROUDIER, SYLVIE
APPLICANT: GIOGI, DOMINIQUE
APPLICANT: KELLY, ROBERT J
TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE
TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE
TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: ORLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,800A
FILING DATE: 28-FEB-1995
CLASSIFICATION: 435
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2115 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 64...1092
US-08-395-800A-7
Query Match
Best Local Similarity 89.9%; Score 83; DB 1; Length 2115;
Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1079 AGAAGCAGCCAGCGGTGTGCGGCACACTGTATCCAGCTACTCGGAGGCTGAGGC 1138
DB 1375 AAAAATATAGCCAGGCGGTGTGCGGCACACTGTATCCAGCTACTCGGAGGCTGAGGC 1434
QY 1139 AGGAGATCGCTTGAACCCGCGGAGCGGAGCTGTGTG 1177
DB 1435 AAGAGATCATCTTGAACCCAGGAGCGGAGGCTGTGAGTG 1473
RESULT 13
US-08-480-784-20
Sequence 20, Application US/08480784
Patent No. 5693473
GENERAL INFORMATION:
APPLICANT: SKOLNICK, MARK H.
APPLICANT: GOLDBERG, DAVID E.
APPLICANT: MIKI, YOSHIO
APPLICANT: SWENSON, JEFF
APPLICANT: KAMB, ALEXANDER
APPLICANT: HARTSMAN, KEITH D.

APPLICANT: Shattuck-Eidens, Donna M.
APPLICANT: Tavligian, Sean V.
APPLICANT: Wiegman, Roger W.
APPLICANT: Futreal, P. Andrew
TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
TITLE OF INVENTION: Susceptibility Gene
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,784
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409,305
FILING DATE: 24-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348,824
FILING DATE: 29-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/308,104
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,266
FILING DATE: 02-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/289,221
FILING DATE: 12-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24884-109347
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 6769 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-480-784-20
Query Match
Best Local Similarity 82.6%; Score 83; DB 1; Length 6769;
Matches 95; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 1079 AGAAGCAGCCAGCGGTGTGCGGCACACTGTATCCAGCTACTCGGAGGCTGAGGC 1138
DB 2396 AAAAATATAGCCAGGCGGTGTGCGGCACACTGTATCCAGCTACTCGGAGGCTGAGGC 2455
QY 1139 AGGAGATCGCTTGAACCCGCGGAGCGGAGCTGTGTGCGGCAGAGCAGCTATCA 1193
DB 2456 AAGAGATCATCTTGAACCCGCGGAGGCTGTGTGCGGCAGAGCAGCTATCA 2510
RESULT 14
US-08-483-553-20

1079 AGAAGCAGCCAGCGCTGTCGCCACACCTGTATATCCAGCTACTGGGAGGCTGAGGC 1138

Query Match 348: Score 83: DB 1: 1anrth 6760

| | Matches | 95; | Conservative | 0; | Mismatches | 20; | Indels | 0; | Gaps | 0; |
|----|---------|--|--------------|----|------------|-----|--------|----|------|----|
| QY | 1079 | ACAAAGCAGCCAGCGCTGTGGCGCACACCTGTATCCAGCTACTCGGAGGCTGAGGC | 1138 | | | | | | | |
| Db | 2396 | AAAAATTAGCCAGGTGTGTGGCACAATACTGTATCCAGCTACTCGGAGGCTGAGGC | 2455 | | | | | | | |
| QY | 1139 | AGGAGATCGCTGTGACCCCGGAGGAGGTGTGTGGCAGAGCAGCTCATCA | 1193 | | | | | | | |
| Db | 2456 | AGGAGATCTACTGTGAACCCCGGAGGTGTGGGTGCGGTGAAACCGAGATCGCACCA | 2510 | | | | | | | |

Search completed: June 8, 2003, 18:16:31
 Job time : 102.098 secs